Cancer prevalence in the United Kingdom
Current estimates, future projections and health service utilisation among cancer survivors

Maddams, Jacob

Awarding institution:
King's College London

Cancer prevalence in the United Kingdom

Current estimates, future projections and health service utilisation among cancer survivors

Jacob Anthony Maddams

King's College London
University of London

2012

Thesis submitted for the degree of PhD

Abstract

Cancer prevalence is an important epidemiological measure of the disease burden. It is defined as the number of people in a given population who are alive at a specified point in time (the index date) and who have previously been diagnosed with cancer. It may be expressed as either a count or as a proportion of the population. Members of the prevalent population are known as 'cancer survivors' and the time spent as such is known as 'cancer survivorship'. *Complete* prevalence includes all survivors regardless of when they were diagnosed, whereas *N-year limited duration* prevalence includes only those who have received at least one cancer diagnosis in the N years prior to the index date.

In the United Kingdom (UK), addressing the needs of cancer survivors is a high priority for the Department of Health, as well as for voluntary sector organisations, and the need for further research into cancer survivorship has been highlighted. Despite this, in recent years little study has focused on cancer prevalence in the UK. The aims in preparing this thesis were to provide up-to-date estimates of cancer prevalence in the UK, to describe levels of acute health service utilisation among cancer survivors in different temporal phases of survivorship and to provide projections of future cancer prevalence.

National cancer registry data for the UK were analysed, together with National Health Service hospital activity data for England. It was found that there are currently around two million cancer survivors in the UK, a figure far higher than previously thought. Levels of acute in-patient health service utilisation were, however, generally low among cancer survivors who had survived at least five years and who were not in the final year of their life. A discrete time model for projecting cancer prevalence was derived and used to project cancer prevalence in the UK from 2009 to 2040 under various different scenarios of future cancer incidence and survival. It was shown that in the coming decades cancer prevalence is likely to increase substantially.

This thesis contains a detailed description of cancer prevalence and aspects of cancer survivorship in the UK which highlights the need for adequate planning to meet the many and varied needs of those diagnosed with cancer.

Contents

List of Figures

List of Tables

Acknowledgements

My sincere thanks go to Professor Henrik Møller and Professor Martin Utley for their invaluable advice, support, inspiration and jokes in the margins; to Macmillan Cancer Support for funding this work; and to Becky Vickers for help with proof reading.

Also to my colleagues at the Thames Cancer Registry, as well as the other UK registries – especially to all those responsible for collecting, compiling and quality assuring the data that underpin this thesis.

Thanks also to everyone in my family for all their love and support over the years.

Finally, most of all I'd like to thank my wife, Alice, whose love, constant support and heroic culinary endeavours have sustained me, who has suffered my obsessive compulsive font-changing disorder with good grace and to whom I owe many more thanks than could be included here without exceeding my word limit!

London, January 2012.

"This is a very complicated case, Maude. You know, a lotta ins, lotta outs, lotta what-have-yous. And, uh, lot of, uh, strands to keep in my head, man, y'know. Lotta strands in old Duder's head."

— The Dude

Chapter 1. Introduction

In this introductory chapter, the basic epidemiological concepts that are used in the rest of this thesis are explained, the motivation for studying cancer prevalence is set out and some background context concerning the United Kingdom (UK) is given. Finally, an overview of the aims of this thesis and its structure is provided.

1.1 Cancer and epidemiology

Cancer is defined by the presence in the body of "an abnormal and unregulated growth of cells [a tumour] which ultimately evolves into a population of cells capable of invading nearby tissues as well as metastasising to distant sites causing significant morbidity and, if untreated, death of the host" (Ruddon, 2007: p.4). In fact, the term 'cancer' includes a rich variety of disease sub-types. Each type is usually distinguished by the location in the body of the primary tumour and by the morphology of the tumour cells. The World Health Organisation's International Classification of Diseases for Oncology (ICD-O) is commonly used to classify the different types (Fritz et al., 2000). Throughout this thesis, cancer types are identified topographically using ICD-10 codes which condense the detail of ICD-O (World Health Organisation, 2010).

Cancer epidemiology is the study of the "distribution and determinants of [cancer]-related states or events in specified populations" (Last, 2001: p.62). Basic epidemiological statistics include incidence, mortality and survival. Cancer incidence is the number of new cancers diagnosed and is generally expressed as a rate – for example, the number of cancers diagnosed per 100,000 population in a given year. Cancer mortality is the number of people who, according to death certification, died from cancer and, similarly, is often expressed as a rate. Cancer survival measures the probability of living a specified amount of time after being diagnosed and can be expressed in a number of ways – for example, as the proportion of cancer patients alive 1, 5 or 10 years after diagnosis. This is known as crude survival. It is also possible to calculate relative survival – an estimate of the probability of surviving in the absence of other non-cancer causes of death (Rutherford et al., in press). Survival is often described as a 'rate' but this is technically inaccurate since a rate should, strictly speaking, be expressed per unit time. Survival is a dimensionless quantity and should correctly be referred to as a probability or, equivalently, a proportion. Therefore, throughout this thesis, the term 'survival rate' is avoided.

Cancer is one of the most significant diseases suffered by humans, is found in almost all other animal species and has been recorded in pictures and writings throughout history (and in recent years, bone cancers have been diagnosed in the remains of ancient Egyptian mummies (Zink et al., 1999)). It is estimated that 12.7 million people worldwide were diagnosed with cancer in 2008 (Ferlay et al., 2010). In the UK, cancer incidence rates have been steadily increasing since records began and, overall, are currently around 500 per 100,000 population, with approximately 300,000 new cases being diagnosed each year (Cancer Research UK, 2011b). It is estimated that, in the UK, one third of people will be diagnosed with cancer during their lifetime and one quarter of all deaths are due to cancer (Cancer Research UK, 2011a).

Unlike many other diseases, systematic recording of the number and types of diagnosed cancers is well established in numerous countries, and the importance of population-based cancer registries to perform this task is widely accepted. In England, there are currently eight regional cancer registries with statutory mandates which, together, cover the whole population of the country. Each is responsible for a separate geographical area, although a national system – as seen in some other European countries, including Wales, Scotland and Northern Ireland – is currently being introduced. The Thames Cancer Registry, for example, covers a population of around 12 million people in South East England and recently celebrated the collation of its 50th year of data (Thames Cancer Registry, 2011). UK cancer registries collect a wide variety of patient and tumour data including demographic, histological, treatment and follow-up information. Generally, tumour registrations are created using a combination of electronic data feeds from hospital and laboratory systems and manual data collection by hospital-based registration officers.

1.2 Prevalence

Cancer prevalence, in addition to cancer incidence, mortality and survival, is another important epidemiological measure. It can be defined in a number of ways, but for the purposes of this thesis 'cancer prevalence' pertains to people who are alive at a given point in time and who have been previously diagnosed with cancer – this is the 'lifetime prevalence' definition (Last, 2001: p.140). It can be expressed as a simple count or as a proportion of the population. Unlike incidence and mortality, it cannot be expressed as a rate since it relates to a static point in time only (the 'index' or 'census' date).

It is possible to calculate *tumour* prevalence – the number of tumours of different types that have been diagnosed in those alive at a given point in time – but it is generally

considered more useful to calculate *person* prevalence, as described above. In this thesis, all analyses pertain to person prevalence, denoted either 'prevalence count' (the number of people) or 'prevalence proportion' (the number of people per 100,000 population, or as a percentage).

In recent times, it has become customary to refer to those people who are included in the cancer prevalence count as 'cancer survivors'. The lifetime definition of cancer prevalence is such that every person is considered to be a cancer survivor from the moment of diagnosis until the moment of death. Clearly this is a broad definition and the population of cancer survivors is therefore highly heterogeneous – from those who are recently diagnosed and in active cancer treatment, to those who have survived long enough to be considered cured. Despite being broad, this definition has many advantages since a diagnosis of cancer has wide-ranging and long-lasting physical, mental and psychosocial consequences for the individual, and disease remission and recurrence may occur many years after the initial diagnosis.

For other diseases, different definitions of prevalence may be more appropriate. For example, when considering AIDS, the prevalence of infection with its causative virus HIV is an important measure since it determines the number of people at risk of developing the disease and the likelihood of a future epidemic (World Health Organisation, 2005). Prevalence of a disease such as seasonal flu, which is caused by different variants of the influenza virus each year, and which, in most cases, has a short duration before recovery, may be defined as the number of people with active disease at a given point in time or as the number of people diagnosed during the previous year. Such definitions lead to the possibility of people moving in and out of the prevalent population over time – something which is not entertained in this thesis when considering cancer prevalence. However, previous work has modelled the progression of cancer patients through different states of disease, from diagnosis and initial treatment through remission, relapse and death (Jackson and Aspden, 1979), with applications to the evaluation of new therapies and clinical trials or options for screening programmes (Jackson et al., 1981, 1982; Jenkins et al., 1994; Gallivan et al., 2007).

The lifetime or 'ever diagnosed' definition of cancer prevalence raises practical considerations and, in this thesis, a distinction is made between 'limited duration prevalence' and 'complete prevalence'. The former includes only those survivors who have been diagnosed with cancer in a given time period prior to the index date – e.g. 5-year prevalence at a particular point in time includes everyone who is alive and has received a diagnosis of cancer in the previous five years. In contrast, complete

prevalence includes every person diagnosed with cancer regardless of how long ago that diagnosis occurred. The best source of data from which to directly estimate cancer prevalence is a population-based cancer registry, but data with an extremely long time series are required to capture complete prevalence, and this is therefore not possible in most cases. The oldest cancer registry in the world is the Connecticut Tumour Registry in the United States which holds 75 years of data (US Department of Public Health, 2011) and from which direct estimates of complete prevalence can be practically made (although, since this is not a national registry, inter-state migration should be considered). In England, quality assured cancer registry data are available for diagnoses made from 1971 onwards, but other registries (for example, the national registries of Northern Ireland and Wales) are younger. Statistical methods must therefore be used to estimate complete prevalence based on limited duration prevalence, and the extent of the available data determines how substantial this adjustment is.

1.3 Motivation

As cancer treatments become more effective and survival increases, the long-term care needs of cancer survivors become more prominent. Increasingly, the prospect of long-term survival or cure is a realistic one for many cancer survivors, leading to a greater focus on long-term management of the disease and its sequelae. In light of this, 'cancer survivorship', defined in broad terms as the experience of living with and beyond cancer, is currently a key focus for the Department of Health in England (as well as for the other organisations responsible for health care in the UK). The National Cancer Survivorship Initiative (NCSI) was established in England in 2008 as a partnership between the Department of Health and the charitable organisation Macmillan Cancer Support. Its primary aim is to "ensure that survivors get the care and support they need to lead as healthy and active a life as possible, for as long as possible" (Department of Health, 2010).

Recent publications in the UK and elsewhere have highlighted the need for a greater focus on cancer survivorship in medical research (Rebbeck et al., 2011; Richards et al., 2011). Such research is key to achieving the vision of the NCSI. Indeed, Richards et al. (2011) suggested that there are 10 specific research questions that need to be answered "if health outcomes for survivors are to be improved". Of particular relevance to this thesis are the following:

- "How many people are currently living with a cancer diagnosis, and how is this likely to change over time?"

- "What care are cancer survivors currently receiving from the NHS in hospitals and in the community?"

Understanding the factors that characterise different phases of cancer survivorship is also a key theme, with questions such as:

- "What specific problems, concerns or needs do cancer survivors report at different times after diagnosis and at different phases in the pathway of care?"
- "What are the risks of survivors experiencing adverse consequences from cancer treatment at different time intervals after diagnosis?"

Evidently, gaining a greater understanding of cancer survivorship and the characteristics of the population of cancer survivors through the study of cancer prevalence is a fundamental starting point for answering these questions. Cancer prevalence pertains to newly diagnosed patients as well as to those at risk of recurrence and/or late effects of treatment and cured patients. Study of cancer prevalence and survivorship therefore has the potential to provide valuable intelligence regarding the characteristics, experiences and needs of *all* people living with cancer. Furthermore, by identifying areas of unmet need, such studies can help to inform the development of appropriate models of care.

From a resource planning perspective, it is important to understand the current and likely future burden of cancer on the health service. This is not only related to the volume of cases and types of *initial* treatment required, but also to the amount of surveillance, rehabilitation, palliative care and treatment for tumour recurrence and secondary and late effects of cancer treatment that is required. Cancer prevalence is one measure of the cancer burden and is complementary to other statistics such as cancer incidence and mortality (Lagiou and Adami, 2002). However cancer prevalence, by considering cancer survivors at all stages of survivorship, provides the most comprehensive description of the cancer burden, especially when augmented with, for example, analyses of health service activity data.

Therefore, the comprehensive account of cancer prevalence and cancer survivorship contained in this thesis forms the basis of much of the intelligence that will be vital in achieving the aims of the NCSI in the coming years, and thus will help, albeit indirectly, to ensure that the best possible health outcomes are achieved for cancer survivors in the UK.

1.4 Aims and outline of thesis

The most recent previous publication to present national estimates of cancer prevalence in the UK came in 2003 from the EUROPREVAL project (Forman et al., 2003), but used a data series ending in 1992. Beyond this, little published research has focused on cancer prevalence in the UK. The aims of this thesis are to provide up-to-date basic estimates of cancer prevalence, to enrich these estimates with information that allows a greater understanding of the different phases of cancer survivorship and the burden of cancer to the individual and to society in general and to provide projections of future cancer prevalence in the UK.

This thesis is broadly arranged into three sections, each addressing one of these aims. In Chapter 2, cancer registry data from all the regional registries in England, as well as the national registries in Northern Ireland, Scotland and Wales, are used to estimate complete cancer prevalence in the UK at the end of 2008. Next, in Chapter 3, a person-time analysis of linked cancer registry and hospital activity data is used to describe the levels of acute health service utilisation among cancer survivors in different temporal phases of survivorship. Finally, the focus of Chapters 4–6 is on projections of cancer prevalence: in Chapter 4 a mathematical model is derived that allows projections of cancer prevalence to be made based on cancer registry data; the required input data for this model is considered in detail and model evaluation exercises are described in Chapter 5; and in Chapter 6 projections of cancer prevalence in the UK up to the year 2040, under various different scenarios of future cancer incidence and survival, are presented.

The scope of the results contained in these chapters is limited to four major groups of cancer types: colon, rectum and anus (ICD-10 C18–C21); lung, bronchus and trachea (ICD-10 C33–C34); prostate (ICD-10 C61); and female breast cancer (ICD-10 C50). These are four of the most commonly diagnosed types of cancer in the UK – together they account for over half of all cancer diagnoses – and exhibit large variations in survival characteristics, from lung cancer with its extremely poor prognosis to breast and prostate cancers with relatively good prognoses. In addition, since it is desirable to provide an overview of the prevalence of all cancers combined, as well as of specific types, analysis is, where appropriate, conducted for a fifth group – 'all other' cancer types – defined as ICD-10 C00–C97 excluding C44 (non-melanoma skin cancer) and the four major types previously mentioned. This group contains a large number of very different types of cancer, from the very rare (such as cancer of the thymus) to relatively common (such as bladder cancer), and is included only to provide a complement to the

other groups from which prevalence of all cancer types combined (excluding non-melanoma skin cancer) can be estimated. Non-melanoma skin cancer is excluded since, historically, it has not been systematically recorded by all cancer registries in the UK.

Attained age (i.e. the age of survivors on the index date) and time since diagnosis are key variables of interest when considering cancer prevalence. In this thesis, attained age is generally grouped into three broad ranges: 0–44, 45–64 and ≥65 years. Where necessary, time since diagnosis is also grouped into three broad ranges: <1, 1–5 and ≥5 years.

Finally, a general summary and discussion of the work is contained in Chapter 7.

1.5 Literature

Searches of relevant literature on the topics of cancer prevalence and survivorship were conducted using standard web-based databases (PubMed, Web of Science, Embase and Scopus). A variety of search terms were used, including "cancer prevalence", "completeness index", "health care" or "quality of life" with "cancer survivors", "cancer survivorship" and "cancer prevalence projections", as well as more specific searches relating to relevant prominent authors. There are three main areas of study in this thesis (estimates of current cancer prevalence, health service utilisation/phases of survivorship and projections of cancer prevalence) and each has a separate body of published research. Therefore the narrative flow of this thesis is maintained by background and literature sections being positioned where they most usefully inform the materials and methods. In section 3.2, the need to go beyond a basic enumeration of cancer survivors when studying cancer survivorship and examples of this in the literature are discussed, including studies of temporal phases of survivorship, cure modelling and health service utilisation. In section 4.2, examples in the literature of approaches to the task of projecting cancer prevalence are presented. This usually involves the description of a mathematical model to relate cancer incidence, survival and prevalence, as well as methods for providing input data to the model, but approaches have often varied depending on the local availability of comprehensive cancer registry data. In section 6.1, previous efforts to assess the independent influences of cancer incidence and survival and population demographics on projections of cancer prevalence are discussed. Relevant literature is also noted in the discussion sections of this thesis (see sections 2.4, 3.5 and 6.5) and elsewhere, in order to place the findings of this work in context.

Chapter 2. Cancer prevalence in 2008

In this chapter, cancer prevalence in the UK at the end of the year 2008 is estimated using national cancer registry data for England, Northern Ireland, Scotland and Wales. Methods are developed to account for cancer survivors who were diagnosed before the earliest available year of cancer registry data, and recent trends in cancer prevalence are described.

This work was first published as an article in the British Journal of Cancer in 2009 (Maddams et al., 2009). The substance and structure of this chapter are therefore the same as that article, but minor edits and amendments to the writing style and figures have been made. A reprint of the original article is contained in Appendix C.

2.1 Introduction

Identifying and addressing the requirements of cancer survivors in England was a high priority in the Cancer Reform Strategy (Department of Health, 2007) and, as a result, the National Cancer Survivorship Initiative was set up in 2008. Similar initiatives are being established in Northern Ireland, Scotland and Wales. However, little is known about the size and demography of the population of cancer survivors in the UK; the most recent estimates of cancer prevalence in the UK, provided by the EUROPREVAL project (Forman et al., 2003), were for 1992.

Cancer survivors may be recently diagnosed and in active treatment, or they may have survived long enough to be considered cured. However, in this analysis, such distinctions are not made; once an individual is diagnosed with cancer, he or she is considered to be a survivor until death. This approach is adopted because a diagnosis of cancer may affect a person's life in different ways (mental health problems, fear of recurrence, financial hardship, relationship issues, etc.), and its effects may be felt for many years after diagnosis. Also, this approach is practical as the currently available cancer registration data do not readily allow survivors to be classified as having active disease, in remission or cured of their cancer.

2.2 Materials and methods

The eight cancer registries in England, together with the national registries in Northern Ireland, Scotland and Wales, provided anonymised records of all registered malignant neoplasms (ICD-10 C00–C97) diagnosed in the residents of those countries, excluding non-melanoma skin cancer (ICD-10 C44) as it is not covered systematically by all

registries. Each record included demographic, tumour, diagnosis, follow-up and death details. Data were available for the periods 1971–2004 for England, 1971–2005 for Scotland, 1990–2006 for Wales and 1993–2006 for Northern Ireland. All tumours apparently diagnosed in patients over the age of 99 years were excluded (approximately 0.04% of the total), leaving 7.7 million registration records for analysis.

The UK cancer registries receive death notifications from the Office for National Statistics (ONS) (England and Wales) and the General Register Offices (Scotland and Northern Ireland), which are then matched to the cancer registration records, although a small percentage are never so matched. The patients associated with these 'lost to follow-up' registrations are, at face value, effectively immortal and result in apparent cancer survivors of a much higher age than is known to be likely. The proportion of registrations lost to follow-up in many European registries is believed to be less than 1% (Micheli et al., 2002), but is not precisely known in UK registries. Therefore, in computing prevalence, cancer survivors were censored at the attained age of 105 years.

Cancer prevalence can be expressed as the number of prevalent tumours or the number of prevalent patients. As each patient may, in their lifetime, be diagnosed with more than one tumour, patient prevalence will always be lower than tumour prevalence. The analysis presented here focused on patient prevalence, and only the first diagnosed malignant neoplasm (excluding non-melanoma skin cancer) in each patient was considered.

Although cancer registry data for Northern Ireland, Scotland and Wales were available for diagnoses made in the years 2005 and 2006, data for England were available only up to 2004 at the time of analysis. For this reason, the most recent index date common to all data was used – i.e. 31 December 2004. The number of people diagnosed with cancer and alive on this date was counted and disaggregated by country of residence, sex, age group on the index date (0–44, 45–64 and ≥65 years), number of years since diagnosis and the following broad groups of cancer diagnoses:

1. Colon, rectum and anus (ICD-10 C18–C21);
2. Lung, bronchus and trachea (ICD-10 C33–C34);
3. Prostate (ICD-10 C61);
4. Female breast (ICD-10 C50);
5. All other malignant neoplasms excluding non-melanoma skin cancer (ICD-10 C00–C97 excluding C44 and (1) to (4)).

Death Certificate Only (DCO) registrations are those for which the only source of patient or tumour information is a death certificate stating cancer as a cause of death. These registrations lack much information, particularly the actual date of diagnosis. An unknown proportion of the DCO registrations since the index date will pertain to patients diagnosed before, and alive on, the index date. However, no attempt has been made to estimate this proportion, and therefore such registrations are not included in the prevalence estimates presented here.

2.2.1 Complete prevalence

N-year limited duration prevalence includes only those survivors diagnosed in the last *N* years before the index date. Complete prevalence includes all cancer survivors, regardless of when they were diagnosed. It is not currently possible to directly count complete prevalence on the basis of cancer registry data alone, given that no UK registry has been collecting data for a sufficiently long period of time. Instead, complete prevalence was estimated by extrapolating from limited duration prevalence as follows.

With an index date of 31 December 2004, the available cancer registry data provided 34-year prevalence estimates for England and Scotland, 15-year estimates for Wales and 12-year estimates for Northern Ireland. To extend these limited duration estimates to complete estimates, a negative binomial regression model with a log-link function was constructed for each type of cancer, sex and age group (0–44, 45–64 and $\geq$65 years). The prevalence count on the index date was the response variable, and the explanatory variables were country of residence and number of years since diagnosis. Given that the primary objective was to obtain a reasonable estimate for the number of people who had survived at least 12 years (Northern Ireland) and 15 years (Wales) beyond diagnosis, data pertaining to the most recent five years of diagnoses were not used in the regression models. Prostate cancer was treated as a special case and modelled in two stages; first, for diagnoses made in the period 1992–1999 and second, for all diagnoses made before 1992. An offset term was also included in all models, defined as the logarithm of the number of people in each country who could contribute to the prevalence count (i.e. the population at risk), taking into account the age group being considered and the fact that years since diagnosis cannot exceed attained age on the index date. The models were run using the `proc genmod` procedure in the SAS statistical programming package (SAS Institute Inc., Cary, NC, USA).

The performance of the regression models was tested by initially excluding from them data for Scotland covering diagnoses made between 1971 and 1992, and by comparing

the modelled estimates with the empirical data for those years. Furthermore, the calculated ratios of 15-year prevalence to complete prevalence (known as the 'completeness index') in England and Scotland were compared with previously published estimates (Forman et al., 2003) – see Table 2.1.

Table 2.1. Comparison of 15-year completeness index, by sex and type of cancer.

	Data source	
	Maddams et al.†	EUROPREVAL‡
Males		
Colon, rectum and anus	0.81	0.87**
Lung, bronchus and trachea	0.58	0.73***
Prostate	0.95	0.97
All malignant neoplasms*	0.78	0.82
Females		
Colon, rectum and anus	0.74	0.80
Lung, bronchus and trachea	0.77	0.79
Breast	0.74	0.80
All malignant neoplasms*	0.70	0.72

The 15-year completeness index is defined as 15-year prevalence divided by complete prevalence. *Excluding non-melanoma skin cancer (ICD-10 C44). †Based on the work described in this chapter: 15-year prevalence divided by estimated complete prevalence in the UK; Index date 31 December 2004. ‡From Forman et al. (2003): Average 15-year completeness index using data from South Thames, West Midlands, Yorkshire and Scotland cancer registries with index date 31 December 1992. **Unweighted average of indices for cancer of the colon and rectum. ***Cancer of the lung only.

2.2.2 Projections from 2004 to 2008

Through an analysis of recent empirical trends in limited duration cancer prevalence, estimates of complete cancer prevalence in the UK were projected from 31 December 2004 forward to 2008. Combined data for England and Scotland (covering cancer diagnoses made in the period 1971–2004) were used to estimate limited duration prevalence between 2000 and 2004, for each sex and cancer type. Log-linear functions, considered appropriate for short-term projections, were fitted to provide estimates of the annual growth in the number of cancer survivors that could be expected from 2004 to 2008. The following assumptions were made:

1. The yearly rates of change of cancer prevalence in England and Scotland combined can reasonably be applied to each constituent country of the UK;
2. The rate of change of cancer prevalence in each age group (0–44, 45–64 and ≥65 years) can be approximated by the rate of change in all age groups combined;

3. For each sex and cancer type, the rate of change of complete prevalence is the same as that of 30-year limited duration prevalence.

Estimated prevalence counts were converted to proportions of the population by using the mid-year ONS population estimates for 2007; these were the most recent estimates available at the time of analysis and likely to be only slightly lower than the actual population at the end of 2008 (Office for National Statistics, 2011a).

2.3 Results

Tables 2.2–2.4 present complete prevalence – the sum of observed prevalence from the years of diagnosis that were available in the data and modelled prevalence from those that were not. Estimates for which more than 20% of the total is derived from modelling are underlined.

It was estimated that at the end of 2008 there were just over two million cancer survivors in the UK – 40.9% were male and 59.1% were female. Approximately 2.7% of the male population, and 3.8% of the female population, were cancer survivors. Table 2.2 shows the variation in cancer prevalence by country, sex and cancer type. Wales had the highest prevalence proportions (3.1% of males and 4.2% of females were cancer survivors), and Northern Ireland had the lowest (2.4% of males and 3.4% of females).

Prostate and female breast cancers were the most prevalent, and accounted for 30.9% of male, and 46.4% of female, cancer prevalence in the UK. Of the cancer types studied here, lung cancer was the least prevalent. Figure 2.1 shows, for each sex, the proportions of total incident cases, cancer deaths and cancer prevalence that were accounted for by colorectal, lung, prostate and female breast cancers (incidence and mortality data taken from the UK Cancer Information Service (National Cancer Intelligence Network, 2011b)). For both males and females, colorectal cancer accounted for approximately 10–15% of all the three measures. In contrast, for males, lung cancer accounted for 15% of all newly diagnosed cancers, 25% of cancer deaths and for only 5% of cancer prevalence. A similar pattern was seen for female lung cancer, which accounted only for 2% of female cancer prevalence. Prostate and female breast cancers provided further contrasts, the latter accounting for 31% of newly diagnosed cancers, 17% of cancer deaths and 46% of cancer prevalence among females.

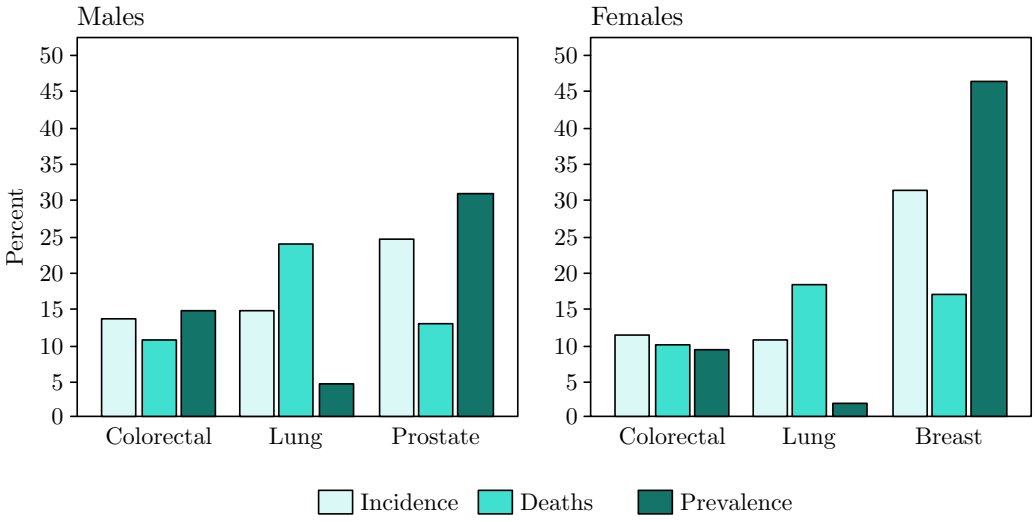
Table 2.3 presents cancer prevalence in the UK, according to time since diagnosis. The proportion of survivors in each time since diagnosis band varied according to sex and cancer type, as illustrated in Figure 2.2. Overall, female cancer survivors tended to be

further beyond their diagnosis than males – 67% of female survivors had been diagnosed more than 5 years earlier, compared with 55% of male survivors.

Table 2.4 shows the variation of cancer prevalence with attained age. Less than 1% of the UK population aged under 45 years at the end of 2008 were cancer survivors, compared with around 13% of those aged at least 65 years. There were twice as many female survivors aged between 45 and 64 years as there were males, largely because of the dominance of female breast cancer that accounted for 54% of female survivors in this age range. The most prevalent types of cancer in those aged at least 65 years were prostate and female breast cancers; 5.2% of males and 5.7% of females in this age group were survivors of these cancers.

Figure 2.3 shows the trends in 1, 5, 10 and 30-year limited duration prevalence in the period 2000–2004, and projected to 2008. Only male lung cancer did not show an increasing trend, the total number of survivors decreasing by 1.4% per year. Prostate cancer prevalence was, by a considerable margin, the most rapidly increasing type – the total number of prostate cancer survivors increased by 9.8% per year in the period 2000–2004. Overall, the number of male cancer survivors increased by 3.8% per year and the number of female cancer survivors increased by 2.7% per year.

Figure 2.1. Proportion of total cancer incidence*, cancer deaths† and cancer prevalence** that is accounted for by selected cancer types‡.



*England, 2006 (data from UKCIS). †England, 2005 (data from UKCIS). **UK, 2008. ‡ICD-10 codes as follows: Colorectal = C18–C21; Lung = C33–C34; Prostate = C61; Breast = C50.

Figure 2.2. Time since diagnosis distribution of cancer survivors in the UK at the end of 2004, by sex and cancer type.

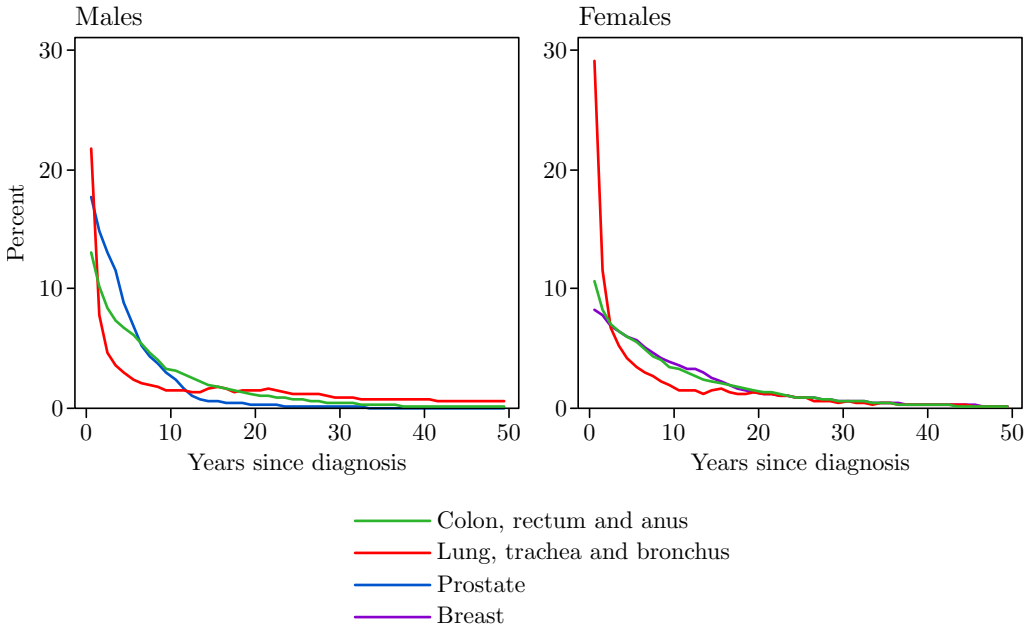
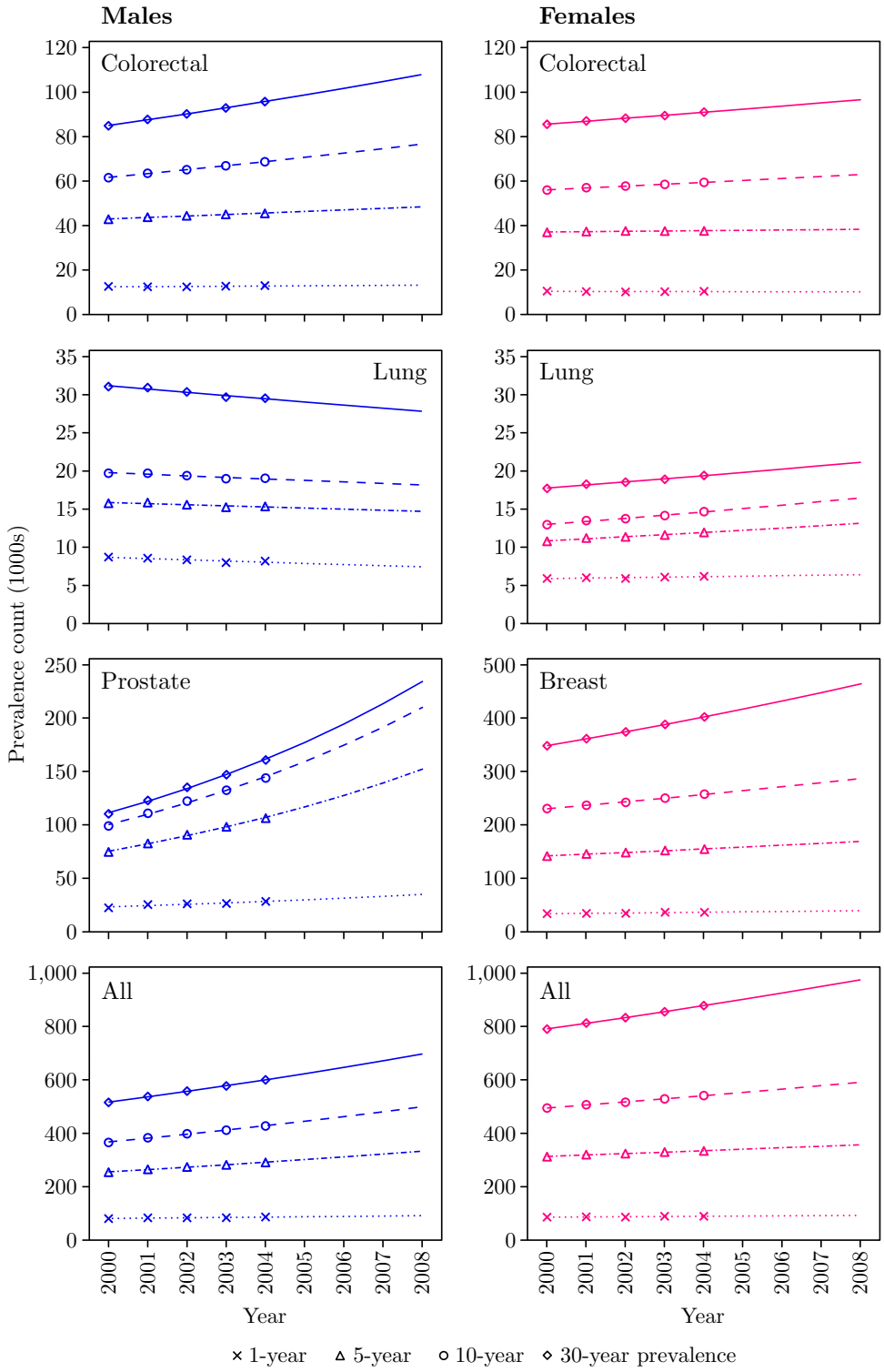


Figure 2.3. Empirical and projected trends in limited duration cancer prevalence in England and Scotland, 2000–2008, by sex and cancer type*.



*ICD-10 codes as follows: Colorectal = C18–C21; Lung = C33–C34; Prostate = C61; Breast = C50. All = C00–C97 excluding non-melanoma skin cancer C44.

Table 2.2. Prevalence of cancer at the end of 2008 in the UK, by country of residence, cancer type and sex. Number of survivors (proportion per 100,000 population).

	Country of residence									
	England		Scotland		Wales		Northern Ireland		UK	
Males										
Colon, rectum and anus	100,608	(401)	11,522	(464)	6,921	(476)	3,480	(404)	122,531	(410)
Lung, bronchus and trachea	32,034	(128)	3,760	(151)	1,889	(130)	1,058	(123)	38,741	(130)
Prostate	215,654	(859)	19,163	(771)	13,312	(916)	5,307	(616)	253,436	(847)
All other malignant neoplasms*	334,147	(1,330)	36,853	(1,483)	22,998	(1,582)	10,482	(1,216)	404,480	(1,352)
All malignant neoplasms*	682,443	(2,717)	71,298	(2,868)	45,120	(3,103)	20,327	(2,358)	819,188	(2,738)
Females										
Colon, rectum and anus	92,439	(356)	11,419	(430)	5,885	(386)	3,542	(395)	113,285	(365)
Lung, bronchus and trachea	19,634	(76)	3,215	(121)	1,239	(81)	693	(77)	24,781	(80)
Breast	460,041	(1,771)	46,211	(1,738)	29,838	(1,955)	12,908	(1,439)	548,998	(1,768)
All other malignant neoplasms*	409,284	(1,576)	46,607	(1,753)	26,683	(1,749)	13,690	(1,526)	496,264	(1,598)
All malignant neoplasms*	981,398	(3,778)	107,452	(4,042)	63,645	(4,171)	30,833	(3,437)	1,183,328	(3,810)

The sum of observed prevalence where available from cancer registry data, and modelled prevalence where not. Underlined numbers are those which are based on estimates of prevalence in 2004 that were at least 20% modelled. *Excluding non-melanoma skin cancer (ICD-10 C44).

Table 2.3. Prevalence of cancer at the end of 2008 in the UK, by time since diagnosis, cancer type and sex. Number of survivors (proportion per 100,000 population).

	Time since diagnosis					
	<1 year	1–5 years	5–10 years	10–20 years	≥20 years	All
Males						
Colon, rectum and anus	14,619 (49)	38,075 (127)	31,162 (104)	24,534 (82)	14,141 (47)	122,531 (410)
Lung, bronchus and trachea	8,263 (28)	7,850 (26)	3,810 (13)	4,769 (16)	14,049† (47)†	38,741 (130)
Prostate	37,967 (127)	125,470 (419)	61,376 (205)	22,601 (76)	6,022 (20)	253,436 (847)
All other malignant neoplasms*	40,891 (137)	97,529 (326)	87,008 (291)	98,726 (330)	80,326 (269)	404,480 (1,352)
All malignant neoplasms*	101,740 (340)	268,924 (899)	183,356 (613)	150,630 (504)	114,538 (383)	819,188 (2,738)
Females						
Colon, rectum and anus	11,309 (36)	30,341 (98)	27,128 (87)	25,532 (82)	18,975 (61)	113,285 (365)
Lung, bronchus and trachea	6,905 (22)	7,255 (23)	3,671 (12)	2,682 (9)	4,268 (14)	24,781 (80)
Breast	42,432 (137)	140,111 (451)	128,672 (414)	145,035 (467)	92,748 (299)	548,998 (1,768)
All other malignant neoplasms*	40,655 (131)	109,179 (352)	96,400 (310)	119,366 (384)	130,664 (421)	496,264 (1,598)
All malignant neoplasms*	101,301 (326)	286,886 (924)	255,871 (824)	292,615 (942)	246,655 (794)	1,183,328 (3,810)

The sum of observed prevalence where available from cancer registry data, and modelled prevalence where not. Underlined numbers are those which are based on estimates of prevalence in 2004 that were at least 20% modelled. *Excluding non-melanoma skin cancer (ICD-10 C44). †Possibly unreliable, see discussion section 2.4.

Table 2.4. Prevalence of cancer at the end of 2008 in the UK, by age, cancer type and sex. Number of survivors (proportion per 100,000 population).

	Age (at the end of 2008)			
	0–44 years	45–64 years	≥65 years	All
Males				
Colon, rectum and anus	2,091 (11)	25,690 (343)	94,750 (2,238)	122,531 (410)
Lung, bronchus and trachea	441 (2)	6,643 (89)	31,657 (748)	38,741 (130)
Prostate	181 (1)	34,511 (461)	218,744 (5,168)	253,436 (847)
All other malignant neoplasms*	68,539 (377)	125,077 (1,671)	210,864 (4,982)	404,480 (1,352)
All malignant neoplasms*	71,252 (392)	191,921 (2,563)	556,015 (13,136)	819,188 (2,738)
Females				
Colon, rectum and anus	2,134 (12)	19,723 (255)	91,428 (1,648)	113,285 (365)
Lung, bronchus and trachea	530 (3)	5,904 (76)	18,347 (331)	24,781 (80)
Breast	25,428 (143)	208,076 (2,694)	315,494 (5,688)	548,998 (1,768)
All other malignant neoplasms*	67,530 (380)	151,756 (1,965)	276,978 (4,994)	496,264 (1,598)
All malignant neoplasms*	95,622 (538)	385,459 (4,990)	702,247 (12,661)	1,183,328 (3,810)

The sum of observed prevalence where available from cancer registry data, and modelled prevalence where not. Underlined numbers are those which are based on estimates of prevalence in 2004 that were at least 20% modelled. *Excluding non-melanoma skin cancer (ICD-10 C44).

2.4 Discussion

The prevalence estimates presented in this chapter were produced using incidence and follow-up data collected by cancer registries in the UK, where available. These estimates were not adjusted to account for DCO registrations that occurred after the index date. DCOs account for less than 5% of all registrations in the UK and most often relate to patients who have died soon after diagnosis. The assumption that their effect on cancer prevalence is small was, therefore, reasonable. No attempt was made to estimate the proportion lost to follow-up, including emigrations. Nor was it possible to include UK immigrants with a diagnosis of cancer pre-dating their move. To a certain extent, the effects of including emigrants and excluding immigrants cancel each other out.

Log-linear regression models of the prevalence count as a function of time since diagnosis were developed and used to estimate the number of survivors from the period before cancer registration in their country. Treating prevalence in this manner as an isolated statistic does not explicitly model the joint effect of incidence and survival, and is based on the observation that the relationship between the number of years since diagnosis and the number of survivors was approximately log-linear in most instances, as in Phillips et al. (2002). However, this relationship was clearly not log-linear for prostate cancer, and so the regression model was, for this cancer, applied in two stages. This was designed to account for the introduction of Prostate Specific Antigen (PSA) testing as a screening tool for prostate cancer in the early 1990s, which effectively changed the definition of the disease with many more localised tumours diagnosed (Evans and Møller, 2003). The number of prostate cancer survivors increased at the fastest rate of the cancers studied here, by almost 10% each year between 2000 and 2004. Since the changes in incidence and survival caused by the introduction of PSA testing are relatively recent, the number of prostate cancer survivors could realistically be expected to increase at a similar rate for some years to come, until a situation is reached in which very few were diagnosed in the era before PSA testing.

Table 2.1 shows that this analysis resulted in 15-year completeness indices for 2004 that were consistently lower than those previously published for 1992 (Forman et al., 2003). A lower completeness index corresponds to a higher proportion of long-term survivors in the prevalent population. The differences in the completeness index are, therefore, consistent with the increases in survival observed between 1992 and 2004 for many cancers (Cancer Research UK, 2009a; Rachet et al., 2008b). Although most differences in these indices were small, for male lung cancer the difference was large (0.58

compared with the published estimate of 0.73). Such a large discrepancy was unexpected, despite the two estimates relating to different index dates and the changes in lung cancer incidence between the two dates. For lung cancer, as with prostate cancer, modelling the number of survivors as a simple log-linear function of time since diagnosis was not appropriate. Owing to its poor prognosis, lung cancer prevalence is dominated by short-term survivors (Figure 2.2) and therefore the regression model has most likely overestimated the number of long-term male lung cancer survivors. If a completeness index of 0.73 was used (as calculated by Forman et al. (2003)), then an estimate of 6,000 (rather than 14,000) would be produced for the number of male lung cancer survivors who, at the end of 2008, had survived more than 20 years. This lower estimate seems more plausible, especially when compared with evidence from other Northern European countries which indicates that in 1992 only 31% of male lung cancer survivors had survived more than 10 years (Möller et al., 2003).

A pragmatic approach to the modelling of cancer prevalence has been adopted in this study. Owing to the long time series of cancer registry data available in the UK, the majority of the estimates presented contain only a small contribution of modelled data. This contribution is most significant in the estimates for Northern Ireland and, to a lesser extent, for Wales. With 34 years of data available for England and Scotland, the modelled proportions of the complete prevalence estimates were typically around 5% for males and 8% for females. Therefore, these estimates are considered to be robust and fit for purpose.

2.4.1 Substantial results of the analysis

There were approximately 2 million cancer survivors in the UK at the end of 2008. Around 13% of all people aged at least 65 years were cancer survivors. Approximately one in three of the UK population will be diagnosed with cancer during their lifetime and one in four will die from it (Cancer Research UK, 2011a); it can also now be stated that around one in eight of those aged at least 65 years are currently cancer survivors.

The overall estimate of 2.0 million cancer survivors at the end of 2008 is far higher than that of 1.2 million at the end of 1992 that was published by Forman et al. (2003). However, this analysis has shown that in recent years the absolute number of cancer survivors in the UK has increased by approximately 3% per annum, and if a similar rate of increase is assumed to apply over the entire period between 1992 and 2008, then the two figures are consistent. Not only is cancer prevalence increasing overall, but the relative prevalence of different types of cancer is also changing. For example, prostate

and female breast cancers have shown some of the largest increases in incidence rates and survival of all cancers in the UK since 1992 (Cancer Research UK, 2009a), resulting in the proportion of total sex-specific prevalence accounted for by each increasing from 14% and 37% in 1992 (Forman et al., 2003) to 31% and 46% in 2008, respectively.

Cancer prevalence varied between the constituent countries of the UK, with Northern Ireland having the lowest prevalence proportion per 100,000 population and Wales the highest. Age-specific proportions for each country are not presented here, but these observed differences are, at least in part, attributable to the different age structures in each country; Northern Ireland has the youngest population (63% aged under 45 years; UK average of 59%) and Wales has the oldest population (18% aged at least 65 years; UK average of 16%). Different patterns of adoption of the PSA test in the early 1990s resulted in higher detection rates of prostate cancer in mainland Britain compared with those in Northern Ireland. Consequently, recorded prostate cancer incidence rates between 1993 and 2003 remained lower in Northern Ireland than in the rest of the UK (Fitzpatrick et al., 2006) and this may also help to explain the lower prevalence of prostate cancer in Northern Ireland.

In areas where cancer registration is less comprehensive than in the UK, models of historical incidence and survival have been developed to estimate current cancer prevalence (Capocaccia and De Angelis, 1997). However, in the work presented here, little modelling was required due to the large amount of available cancer registry data and for simplicity prevalence was treated as an isolated statistic. Nevertheless, it is important to appreciate that prevalence is not a completely isolated measure, and that historical incidence and survival combine to produce the prevalence figures of today. Figure 2.1 provides an illustration of this: a cancer with a poor prognosis, such as lung cancer, accounts for a very small proportion of prevalence, despite being one of the most commonly diagnosed cancers. Conversely, prostate and female breast cancers, with relatively good prognoses, account for larger proportions of prevalence than they do for new incident cases. Changes in incidence and survival (brought about by changes in lifestyle, population demographics, cancer diagnosis and treatment, health service policy, etc.) will therefore have significant consequences for the prevalence of cancer. For example, since the 1970s, the proportion of the UK male population who are smokers has decreased from over 50% to under 30% (Davy, 2006), resulting in a decrease in male lung cancer incidence and, in turn, a decrease in prevalence. The

effects of changing incidence rates, survival and population demographics on cancer prevalence are explored in greater depth in Chapter 6.

The previous most recent estimates of UK cancer prevalence related to 1992 (Forman et al., 2003). Since then, cancer prevalence has changed markedly. Therefore, the up-to-date estimates contained here are highly relevant for both statutory and voluntary sector organisations that are responsible for planning and providing treatment and support to cancer survivors in the UK. In the coming years, cancer prevalence is likely to continue increasing as a result of the growing and ageing population of the UK, increased detection of cancer and longer survival. It was estimated that, overall, the annual rate of increase in the number of cancer survivors is currently around 3%, and it is anticipated that this rate of increase will continue in the near future. Issues surrounding care and support for cancer survivors ought to, therefore, remain high on the public health agenda, and analysis and projections of cancer incidence, prevalence and mortality should become increasingly central to resource planning decisions. Cancer detection and treatment resources tend to focus on the most commonly diagnosed types of cancer, or those that cause the most deaths, but when considering cancer prevalence the most significant cancers are those with both a high incidence rate and a relatively good prognosis in terms of survival (such as prostate cancer and female breast cancer).

Knowledge of the natural progression of a particular type of cancer, together with analysis of prevalence according to time since diagnosis, gives some indication of different phases of survivorship, but does not fully show the extent to which survivors require, or are receiving, care and support. Many survivors will be newly diagnosed and in active treatment, others may be in a state of remission or recurrence with or without late effects of treatment, others may be receiving palliative care, whereas some may consider themselves to be completely free of cancer. Awareness of these differences is important when assessing the health care burden of cancer, the financial costs of which have been shown to be highest during initial treatment and end of life care (Brown et al., 1999). The work presented in this chapter is extended to provide additional details about the different phases of survivorship (see Chapter 3), as well as to analyse the factors that influence cancer prevalence over time (see Chapter 6).

Chapter 3. Health service utilisation and phases of survivorship

In Chapter 2, estimates of cancer prevalence in the UK at the end of 2008 were presented for a variety of different cancer types, according to attained age and time since diagnosis. These are significant as the most up-to-date national estimates of cancer prevalence in the UK.

In this chapter, the basic description of cancer prevalence in the UK is enhanced with information about the level of health service utilisation among cancer survivors in different temporal 'phases' of survivorship. Much of the work presented in this chapter was published in 2011 as two journal articles (Maddams et al., 2011a, b) – reprints of these articles can be found in Appendix C.

3.1 Introduction

There are approximately two million cancer survivors in the UK and in recent years this number has increased by approximately 3% per annum (Chapter 2). The increasing prevalence of cancer in the UK is largely due to the ageing population, earlier detection of cancers and improved treatment regimes leading to increased recorded survival. The Cancer Reform Strategy, (Department of Health, 2007) highlighted the need for a National Cancer Survivorship Initiative (eventually established in 2008) to focus on the needs of the growing population of survivors. Cancer survivorship has been defined as the experience of 'living with or beyond cancer' – the time spent as a survivor – and is often described in such terms (McCabe, 2007; Cooley, 2010a, b; Nyatanga, 2010; National Cancer Survivorship Initiative, 2011). Yet, until recently, there has been little co-ordinated study of cancer survivors' experiences and interaction with health services.

It is important that a study of cancer survivorship adds to our understanding of the cancer burden in terms of a) the personal psychosocial and physical burden to the individual survivor; and b) the resource burden to the health service and society at large. Whilst valuable, a simple enumeration of cancer survivors tells us only a limited amount about the burden presented by the disease in these terms, and there is therefore a need for more detailed analyses of cancer prevalence and cancer survivors in the UK.

In this chapter, person-time analyses of a linked cancer registry and health service activity dataset are presented. Firstly, acute health service utilisation among cancer survivors in the UK is described according to attained age and time since diagnosis.

Secondly, four different 'phases of survivorship' are defined in terms of time since diagnosis and time until death. These phases are then combined with three categories of health service utilisation intensity, as well as three broad age groups, to define a set of survivorship states. The number of cancer survivors in the UK at the end of 2008 in each of these states is then estimated.

3.2 Background and literature

Recorded cancer survival in the UK has increased substantially over the last 25 years (Rachet et al., 2008b, 2009). Large increases have been observed particularly for colorectal, female breast and prostate cancers, which have seen deprivation adjusted increases in 10-year relative survival of around 5%, 10% and 16%, respectively (Mitry et al., 2008a, b; Quinn et al., 2008; Rowan et al., 2008). Lung cancer, although one of the most commonly diagnosed cancers in the UK, continues to have extremely low survival which has remained largely constant since the 1970s (Rachet et al., 2008a). Earlier detection of cancer through, for example, the PSA test for prostate cancer, mammography screening programmes for female breast cancer and better public awareness of early cancer symptoms, as well as improved treatment regimes, have all contributed to the observed increases in survival in the UK (partly due to the introduction of lead time or length bias (Duffy et al., 2008)). Consequently, for many people diagnosed with cancer it is no longer considered the death sentence it once was (Maliski et al., 2002; Hubbard, 2010). Indeed, it has been argued that in many cases cancer may be better described as a chronic illness, i.e. one characterised by a prolonged duration and a recurring nature (Markman, 2006). For some, this description is unhelpful (Vera-Garcia, 2005), but for others the proposed shift towards thinking of cancer as a chronic illness is positive and provides the opportunity for long-term planning of effective disease surveillance and intervention (Phillips and Currow, 2010).

As survival increases, so does the number of long-term survivors and survivorship experiences become more diverse. The study of survivorship must therefore go beyond a simple enumeration of survivors. Describing cancer prevalence according to pre-defined phases of survivorship can provide a greater understanding of the population of survivors, their needs and experiences. Prevalence is often estimated directly from cancer registry data, and using these data it is convenient to simply disaggregate the prevalence estimates according to year of diagnosis or equivalently time since diagnosis (Capocaccia et al., 1997; Benhamiche-Bouvier et al., 2000; Parkin et al., 2001; McCarthy, 2002; Pisani et al., 2002; Gatta et al., 2004; Louchini et al., 2006; Yabroff et

al., 2008a). Intuitively this appears a good way to broadly categorise cancer survivors – certainly the first year following diagnosis is likely to be one of the most traumatic physically (as initial treatment is received) and emotionally (as survivors adjust to life post-diagnosis). However, in the medium and long term after diagnosis, survivorship experiences are likely to vary greatly across individuals, and it may not be sufficient to classify survivors by time since diagnosis alone.

Dates of death are also available in data from most cancer registries and can be used to identify, retrospectively, those survivors who are nearing the end of their life – for example, those who are less than one year from death. When combined, time since diagnosis and time until death can be used to define more detailed temporal phases of survivorship, such as 'initial treatment', 'follow-up and monitoring' and 'end of life' (Brown et al., 1999; Mariotto et al., 2006).

Survival analysis has also been used to classify survivors. Cure modelling suggests that, for certain types of cancer, it is possible to estimate the proportion of patients falling into each of two separate groups: those bound to die of their cancer, and those who will eventually be 'cured' in the sense that, after a period of time, they have no excess mortality risk when compared with the general population (Coldman et al., 1992). Cured proportions can therefore be defined for cancers with relative survival functions that are approximately constant a certain amount of time after diagnosis. Survivors of colorectal cancer have been shown to suffer little excess mortality compared with the general population after about 6–8 years of survival, and therefore could be considered 'cured' (Verdecchia et al., 1998; Lambert et al., 2007; Smastuen et al., 2008), but this is not possible for all cancers – for example, survivors of female breast cancer are unlikely ever to reach a point of cure in this sense (Brenner and Hakulinen, 2004; Francisci et al., 2009). Nonetheless, statistical cure defined in this fashion has been used to estimate the number of 'cured survivors' who could be considered as no longer contributing to the cancer burden (Capocaccia et al., 1997; Verdecchia et al., 1998; Phillips et al., 2002; Gatta et al., 2004; Lambert et al., 2007; Francisci et al., 2009). This is a rather heavy-handed approach to survivorship though, since the concept of statistical cure is quite theoretical, and distinct from clinical cure. Indeed, even a genuinely, clinically 'cured' survivor may continue to suffer the consequences of their diagnosis through late effects or disability from treatment, financial hardship, anxiety and depression, fear of recurrence, etc. (Simonelli et al., 2008).

Another dimension that can augment the basic prevalence statistics is the use of health care services by cancer survivors. This not only helps to describe one aspect of the

survivorship experience, but also provides an estimate of a proportion of the financial and resource burden of the disease to health care services and society. Detailed information regarding health service utilisation is not generally available in cancer registry data, and so an external data source is required as a supplement. In the UK, hospital activity data are routinely collected and have been used in Scotland to analyse hospital admission rates in terms of the number of episodes of care (not specific to cancer or cancer survivors) per 100,000 population (Information Services Division Scotland, 2011), as well as the average lengths of stay and number of bed days utilised. In the USA, the SEER-Medicare linked dataset has been used to estimate the burden of colorectal cancer by analysing the quantity of health care given to survivors in four phases of survivorship – initial diagnosis and treatment, post-diagnostic monitoring, treatment for recurrent/metastatic disease and terminal care (Mariotto et al., 2003). The patterns of recurrence or later metastatic disease in cancer survivors have also been the subject of direct cohort follow-up studies in Europe (Benhamiche-Bouvier et al., 2000; Colonna et al., 2001; Gatta et al., 2004), and these have led to estimates of the proportion of survivors who are in complete remission – a complement to the estimates of statistically cured survivors. In France, Colonna et al. (2012) demonstrated how a joint analysis of cancer registry data and hospital medico-administrative data could be used to estimate 'hospital prevalence' – i.e. the number (and proportion) of cancer survivors admitted to hospital for cancer-related care within a given time period. They found that, of the women who had been diagnosed with breast cancer in the previous 33 years, 17% were hospitalised during 2007. This was considered to be a good measure of the amount of care required by cancer survivors, although explicit consideration was not given to any unmet needs.

In certain countries it has been possible to use hospital activity and administrative data to directly estimate the financial costs associated with cancer at different stages of the survivorship pathway. Yabroff et al. (2008b) developed the work of Mariotto et al. (2003) by estimating the financial costs incurred by Medicare (the medical insurance company) relating to treatment and care of colorectal cancer survivors in the USA. By matching cancer patients to non-cancer controls in the dataset, cancer-specific costs were estimated for the 'initial', 'continuing' and 'last year of life' phases of survivorship, and projected from the year 2000 to 2020. The estimated costs of colorectal cancer care among those aged over 65 were projected to increase substantially – by 53% over the 20 year period if incidence, survival and financial costs remained constant, and by 89% if existing trends continued. A similar analysis of the SEER-Medicare database was

conducted by Ramsey et al. (2002) to estimate the lifetime cancer-attributable costs of care for long-term colorectal cancer survivors. The financial cost of medical care for these cancer survivors was found to be significantly higher than matched non-cancer controls, even up to 10 years after the date of diagnosis. In Sweden, Norlund et al. (2003) used cancer registry and regional administrative data to estimate the total costs of in-patient and out-patient healthcare for prostate cancer survivors up to three years after diagnosis. They found that costs were highest in the first year after diagnosis and decreased in subsequent years.

It is clear that as the cancer care agenda in the UK shifts attention towards survivorship, there is a need for more in-depth prevalence statistics that identify distinct sub-groups of the cancer survivor population and describe the experience of survivorship in more detail.

3.3 Materials and methods

3.3.1 Data

The analysis presented in this chapter was based on two datasets which were linked at the patient level: the English national merged cancer registry dataset and the English Cancer Registries' National Hospital Episode Statistics (HES) extract. The former featured patient and diagnostic information relating to all cancers diagnosed between 1990 and 2006 and recorded by the eight regional population-based cancer registries in England which, together, provide 100% geographical coverage of the country. HES is a record level data repository managed by the National Health Service Information Centre on behalf of the Secretary of State for Health. It contains patient, clinical and administrative details for admitted patients and out-patients treated in any hospital operated by the National Health Service (NHS) in England, and is mainly populated by extracts from routine data flows exchanged between health care providers and commissioners (NHS Health and Social Care Information Centre, 2010). Each HES record defines a Finished Consultant Episode (FCE) of care under a given consultant in a given NHS provider. A patient's journey from admission to discharge may be made up of many FCEs. The English Cancer Registries' National HES extract is a subset of the complete HES database and contains only episodes for admitted patients (i.e. in-patients and day case patients) who have at least one recorded episode 'for or with' cancer. An episode is considered to be 'for or with' cancer if any of its 14 diagnostic fields contain an ICD-10 code between C00 and C97 (malignant neoplasms), between

D01 and D48 (in situ, benign or uncertain neoplasms) or equal to O01 (Hydatidiform mole).

The linkage between these two datasets was designed and developed jointly by the Thames Cancer Registry and the Northern and Yorkshire Cancer Registry and Information Service on behalf of the UK Association of Cancer Registries and the National Cancer Intelligence Network. The methodology was 'rules-based' and used NHS number (which is unique to each patient), date of birth, date of death where appropriate, sex and postcode of residence to match HES episodes of care to cancer patients in the national cancer registry dataset.

The national cancer registry dataset was used to define a cohort of cancer survivors who had been diagnosed with a malignant neoplasm (ICD-10 C00–C97 excluding C44) in the period 1990–2006, and alive for at least some portion of 2006. Sub-cohorts were defined according to type of cancer: colon, rectum and anus cancers (ICD-10 C18–C21); lung, bronchus and trachea cancers (ICD-10 C33–C34); prostate cancer (ICD-10 C61); and female breast cancer (ICD-10 C50). Survivors with multiple diagnoses were permitted to be members of multiple sub-cohorts. The unique patient identifiers produced by the linkage algorithm were then used to extract all HES episodes of care that occurred in, or overlapped, the year 2006 for the cancer survivors in each cohort.

Not all of the extracted episodes mentioned cancer, since the English Cancer Registries' HES extract contains *all* in-patient and day case episodes (from any time period) for those patients with at least one cancer related episode. Episodes not mentioning cancer may pertain to an entirely unrelated condition, or (less likely) may have been incorrectly coded. Equally, some cancer survivors had no matching HES episode of care in 2006, due to either a failure to register episodes that did occur, a failure in the matching procedure between the two datasets or simply because they were not admitted to hospital in 2006. In this analysis, matched episodes of care were considered to be 'cancer related' if one of the 14 diagnostic codes was between C00 and C97 (excluding C44), otherwise they were considered to be 'non-cancer related'.

3.3.2 Hospital activity among cancer survivors

Traditionally, prevalence estimates are calculated as a static enumeration of survivors at a given point in time, for example the end of the year. An alternative formulation is to count the amount of person-time that a population spends in cancer survivorship during a given period of time, for example a complete calendar year. When considering the interaction of survivors with the health service, the person-time formulation has a

distinct advantage since it allows us to estimate the proportion of survivor-time spent admitted to hospital as well as to count the number of admissions. For this reason, a person-time approach was used throughout this analysis.

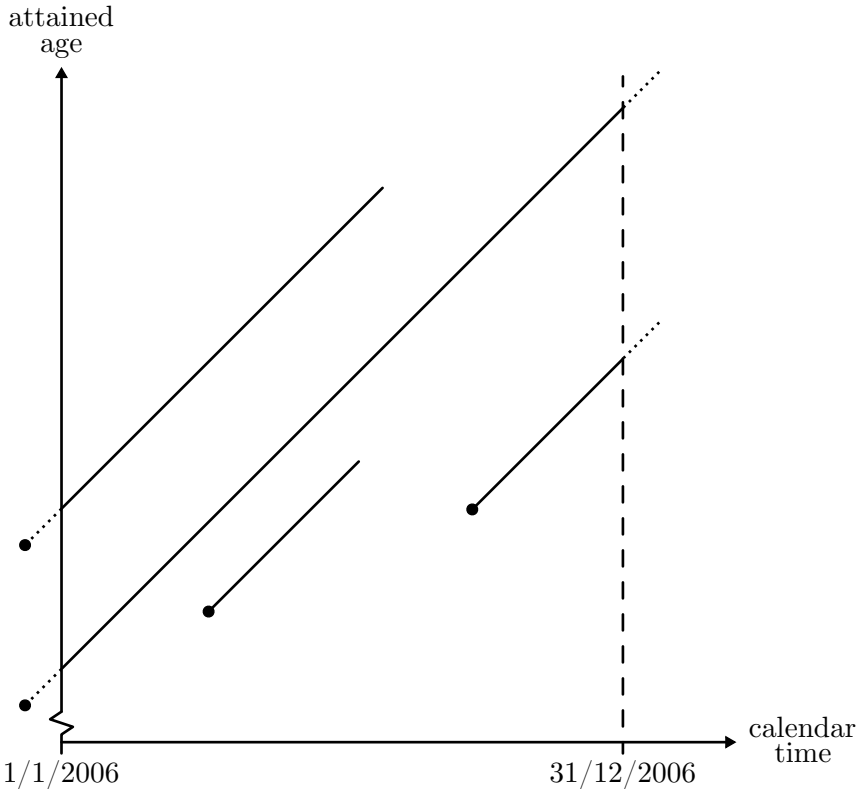
The first goal of the person-time analysis was to describe the effect of attained age and time since diagnosis on the amount of hospital activity in the population of cancer survivors. The most recent calendar year for which data were available, 2006, was chosen to be the period of analysis. Attained age was considered in 5-year groups up to and including 84 with a final group for those aged 85 and over. Time since diagnosis was considered in the broad ranges: <1, 1–5 and ≥ 5 years. Diagnoses were available for the period 1990–2006 inclusive, so in the data there were some survivors prevalent during 2006 who had been diagnosed more than 16 (but less than 17) years previously. However, the 2006 cohort did not contain all such survivors (since no diagnoses from 1989 were available) and therefore the maximum time since diagnosis considered was 16 years. Hence, the '≥ 5 years' time since diagnosis band actually related to 5–16 years since diagnosis.

For each survivor in each cohort, the person-time for which they were prevalent in 2006 (i.e. that which was post-diagnosis and pre-death and overlapped the calendar year 2006) was split into segments according to the time points at which the indexing variables 'attained age' and 'time since diagnosis' changed. This was achieved using the SAS statistical software package (SAS Institute Inc., Cary, NC, USA) and a series of programs developed specifically for this task, based on the program `Lexis.sas` by Carstensen (2007). This approach is similar to that used when conducting survival analyses using the 'period' method, where cross-sectional person-time at risk is considered for a given calendar period (Brenner and Gefeller, 1997).

A Lexis diagram (Lexis, 1875) can be used to illustrate this process. This diagram – named after the German demographer Wilhelm Lexis (1837–1914) and an example of Stigler's Law of Eponymy (Stigler, 1999; Vandeschrick, 2001) – is a simple chart that can be used to present population dynamics according to three demographic co-ordinates: cohort of birth, calendar time and attained age. These co-ordinates are linearly dependent (since cohort of birth is defined as calendar time minus attained age) and can therefore be represented on a two-dimensional chart. Lexis originally proposed a (cohort, age) co-ordinate system, but modern Lexis diagrams use a (calendar time, age) co-ordinate system with each person represented by a line segment starting at birth and ending at death, in such a way that the slope of the line is 1 (Keiding, 2006). A single trajectory in the Lexis diagram may be shared by more than one individual in

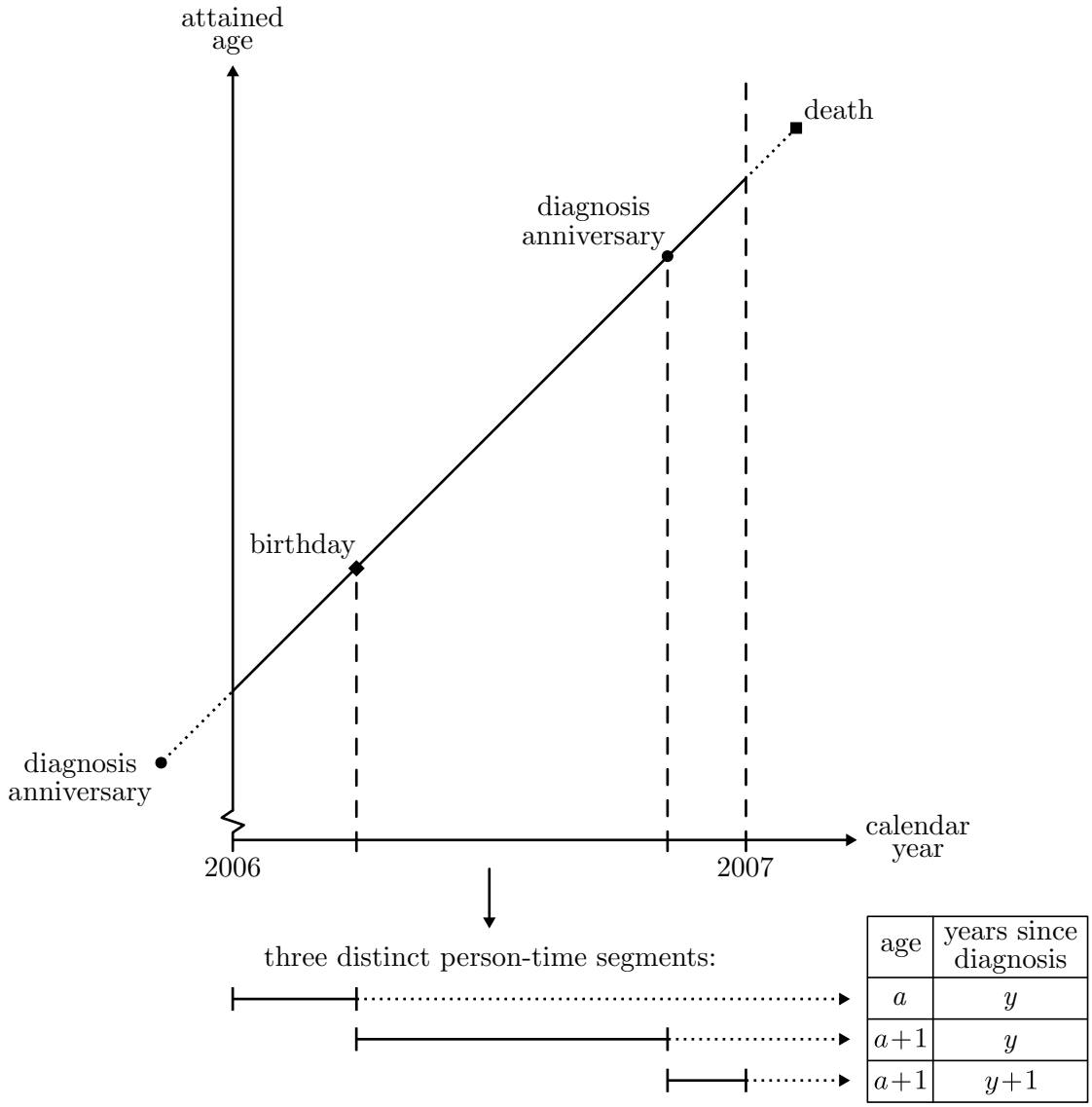
a population. Figure 3.1 depicts a Lexis diagram for four cancer survivors prevalent during 2006, with the line segments starting at the date of diagnosis (rather than birth) and ending, where applicable, at death.

Figure 3.1. Example Lexis diagram for cancer survivors prevalent during 2006.



During the calendar year, each survivor has two possible anniversaries – the points at which either their discrete attained age or time since diagnosis in years increases by one. These anniversaries were used to classify their person-time of prevalence (see Figure 3.2). By summing the lengths of the person-time segments along the calendar time axis in the Lexis diagram, the total amount of person-time of prevalence in 2006 for each cohort of survivors was calculated according to 5-year attained age group and single year since diagnosis.

Figure 3.2. Illustrative survivor person-time segments for the year 2006, indexed according to attained age and time since diagnosis.



Using similar methods, the person-time of hospital activity among cancer survivors was also calculated according to 5-year attained age group and single year since diagnosis. Day case episodes with the same start date and end date were considered to have a duration of one day – therefore, for consistency, the duration of an episode featuring overnight stays in hospital was calculated as the number of days between the start date and end date, inclusive. The proportion of total person-time each cohort of survivors spent admitted to hospital, according to attained age and time since diagnosis, was then calculated by dividing the total population person-time of hospital activity by the total population person-time of prevalence. This quantity may be interpreted as the mean proportion of time in the analysis period spent admitted to hospital by a survivor

of a given age and time since diagnosis sampled at random. Or, equivalently, as an estimate of the probability of that survivor being an admitted hospital patient on any given day in the period. It may be presented as the number of days spent admitted to hospital per 100 person-days, or as a percentage.

To formalise these methods, the following definitions are made for a cohort of N survivors prevalent for some portion of the time period $(0, T]$ (for example, a calendar year) where each individual is indexed by i, with $i \in [1, N]$:

$$X_{i,a,y,t} = \begin{cases} 1 & \text{if individual } i \text{ is a survivor aged } a \text{ and } y \text{ years} \\ & \text{post-diagnosis at discrete time } t \in [1, T] \\ 0 & \text{otherwise} \end{cases} \quad [3.1]$$

$$H_{i,a,y,t} = \begin{cases} 1 & \text{if individual } i \text{ is a survivor and an admitted hospital patient} \\ & \text{aged } a \text{ and } y \text{ years post-diagnosis at discrete time } t \in [1, T] \\ 0 & \text{otherwise} \end{cases} \quad [3.2]$$

where y is measured discretely as the number of whole years that have passed since diagnosis.

Then, the number of time units (e.g. days) spent as a survivor of attained age a and y years since diagnosis by individual i in the time period $(0, T]$ is given by:

$$\sum_{t=1}^{T} X_{i,a,y,t} \, . \quad [3.3]$$

Therefore, summed over all individuals in the population, the total person-time of prevalence at attained age a and y years since diagnosis is given by:

$$S_{a,y} = \sum_{i=1}^{N} \sum_{t=1}^{T} X_{i,a,y,t} \, . \quad [3.4]$$

Similarly, the total person-time of hospital activity among survivors aged a and y years since diagnosis is given by:

$$J_{a,y} = \sum_{i=1}^{N} \sum_{t=1}^{T} H_{i,a,y,t} \, . \quad [3.5]$$

Then, the proportion of all time spent admitted to hospital by survivors aged a and y years since diagnosis is given by:

$$R_{a,y} = \frac{J_{a,y}}{S_{a,y}} \,. \qquad [3.6]$$

This is also the mean proportion of time in the period $(0, T]$ spent admitted to hospital for an individual sampled at random from the prevalent population.

Suppose for a given survivor i, aged a and y years since diagnosis, that $X_{i,a,y,t} = 1$ for all $t \in [t_1, t_2]$, and $X_{i,a,y,t} = 0$ for all $t \notin [t_1, t_2]$. In other words, the person-time of prevalence segment starts at t_1 and ends at t_2. Let $m_{i,a,y} = t_2 - t_1 + 1$ be the duration (in days) of the person-time segment. Then, the random variables $\{H_{i,a,y,t} : t \in [t_1, t_2]\}$ form a sequence of $m_{i,a,y}$ Bernoulli trials and each follows the binomial distribution. However, these are not independent trials since days in hospital for an individual survivor are likely to be clustered together in episodes or spells of care, rather than randomly distributed in time. This situation is called positive serial dependence (Budescu, 1985) and must be accounted for when estimating confidence intervals for $R_{a,y}$.

Perhaps the simplest way to model this dependence is to assume that the random variables $\{H_{i,a,y,t} : t \in [t_1, t_2]\}$ form a stationary first order Markov chain – i.e. the probability that $H_{i,a,y,t} = 1$ is independent of all preceding trials with the exception of that which immediately precedes it (the first order Markov condition), and this probability is independent of the trial number (the stationary condition) (Feller, 1971: p.94–99). The standard binomial model for independent trials is modified for the dependent situation by including an additional dependence parameter, $\lambda = \Pr(H_{i,a,y,t} = 1 \mid H_{i,a,y,t-1} = 1)$ for $t \in [t_1 + 1, t_2]$ (Klotz, 1973). Budescu (1985) gave a method for estimating λ using maximum likelihood estimates of the transition probabilities between consecutive Bernoulli trials, and provided an approximate confidence interval for the success probability (i.e. in this context, the probability of a day being spent in hospital) when estimated from a sample of N individuals, where dependence exists within each subject's trials but there is independence between subjects. This method was used to estimate confidence intervals for $R_{a,y}$, the proportion of time spent in hospital activity by cancer survivors.

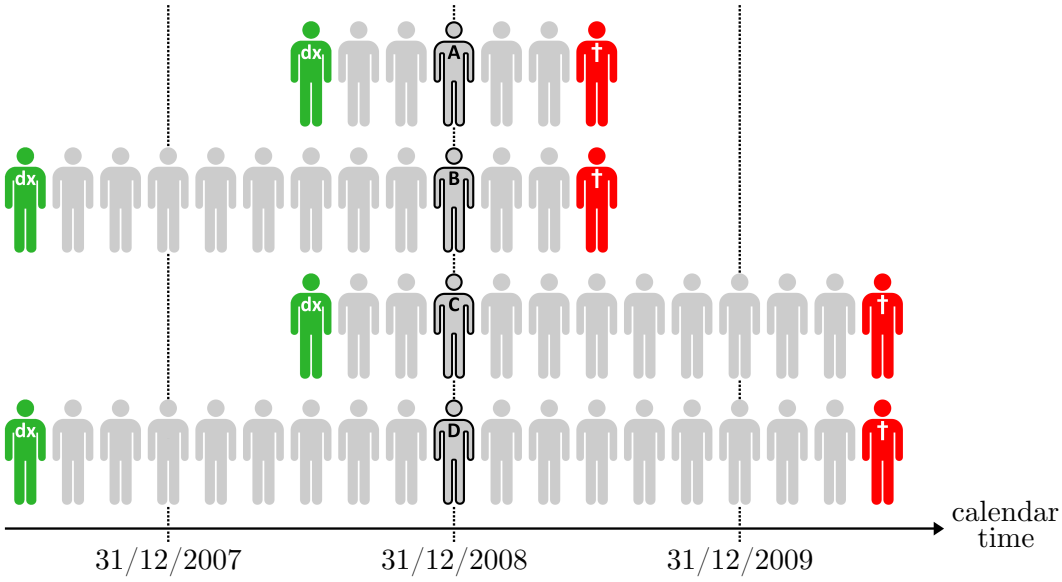
3.3.3 Survivorship states

To provide a means for further enriching the basic prevalence statistics, distinct temporal phases of survivorship were defined according to time since diagnosis and time until death. Specifically, at a given point in time, distinctions were made between survivors who were in the first year after diagnosis and those who were not, and

survivors who were in the last year of their life and those who were not. Four temporal phases of survivorship, depicted in Figure 3.3, were thus defined:

- Phase A: less than one year before death and less than one year after diagnosis.
- Phase B: less than one year before death and more than one year after diagnosis.
- Phase C: more than one year before death and less than one year after diagnosis.
- Phase D: more than one year before death and more than one year after diagnosis.

Figure 3.3. Examples of survivors in temporal phases A–D on 31 December 2008.



An analysis of these temporal phases requires dates of diagnosis and dates of death (data readily available in UK cancer registries) and, at a given point in time, every survivor will be in one, and only one, phase. They therefore provide a practical and useful way of sub-dividing the population of cancer survivors, and are largely comparable to the definitions used by Brown et al. (1999) and Mariotto et al. (2006) who both defined the 'initial' phase according to time since diagnosis ($\leq$6 or $\leq$12 months), the 'end of life/terminal' phase as the last year of life and the 'continuing care/monitoring' phase as everything in between.

3.3.3.1 Analysis for England in 2006

Details of cancer diagnoses among residents of England made in the period 1990–2006, and the corresponding hospital activity data for these patients, were available in the linked cancer registry and hospital activity dataset. Death notifications (provided in England by the Office for National Statistics) were available for survivors in the cancer registry dataset up to the end of 2008, so it was possible to identify the points in time during 2006 at which any survivors entering the last year of their life did so. For each survivor, the person-time for which they were prevalent in 2006 (i.e. that which was post-diagnosis and pre-death and overlapped the calendar year 2006) was split into segments according to the time points at which they moved between temporal phases or broad age groups (0–44, 45–64 and ≥65 years).

For each survivor, each segment of person-time was then indexed according to 'intensity of acute health service utilisation' by calculating the combined duration of cancer related hospital episodes that occurred within it. 'Intensity of acute health service utilisation' was considered to be 'high' if hospital activity accounted for more than 10% of the person-time segment; it was considered to be 'low' if hospital activity accounted for some, but no more than 10%, of the segment; and a separate category, 'none', was reserved for those segments of person-time which contained no hospital activity.

There were, therefore, 12 possible survivorship states defined according to temporal phase and intensity of acute health service utilisation (Table 3.1). The total amount of person-time spent by the population of cancer survivors in each of these states was calculated separately for each cancer type, sex and broad attained age group.

Table 3.1. Twelve survivorship states.

Temporal phase	Intensity of acute health service utilisation
A	None
A	Low
A	High
B	None
B	Low
B	High
C	None
C	Low
C	High
D	None
D	Low
D	High

Under certain assumptions, the proportion of total population person-time spent in each survivorship state is equal to the proportion of the prevalent population in that state at a given point in time during the analysis period (see Lemma 3.1 below). Using this fact, the summed population person-time described above was used to estimate the proportion of survivors in each state at the end of 2006 in England.

Lemma 3.1

Consider a population of survivors, $\{i\}$, each of whose person-time of prevalence overlaps the time interval $(0, T]$. The actual number of survivors will fluctuate during this period. But suppose that the number and attained age of survivors at a given point in time in the period, say $u \in (0, T]$, is known, as is the total amount of person-time spent in each phase of survivorship and age group during the period.

Let $t \in (0, T]$ be the discrete time sampling point at which prevalence is evaluated;

Let $N_{a,t}$ be the number of survivors aged a at time t;

Let $n_{a,p,t}$ be the number of survivors aged a in phase of survivorship p at time t;

Let

$$X_{i,a,p,t} = \begin{cases} 1 & \text{if survivor } i \text{ is aged } a \text{ and in phase } p \text{ at time } t \\ 0 & \text{otherwise} \end{cases};$$

Let S_a be the total person-time of prevalence at age a, summed over all survivors in the period $0 < t \leq T$;

Let $s_{a,p}$ be the total person-time of prevalence spent in phase p at age a, summed over all survivors in the period $0 < t \leq T$;

Then, if we assume that the distribution of $N_{a,t}$ is in steady state throughout the period $0 < t \leq T$, then the expected number of survivors aged a in phase p at time t sampled at random from the interval $(0, T]$ is estimated by the total number of survivors at time T multiplied by the proportion of total person-time of prevalence at age a that is spent in phase p during the period.

Proof

$$n_{a,p,t} = \sum_i X_{i,a,p,t} \quad \text{and} \quad s_{a,p} = \sum_i \sum_{t=1}^{T} X_{i,a,p,t} \tag{3.7}$$

$$\begin{aligned} \Rightarrow \bar{n}_{a,p,t} &= \sum_{t=1}^{T} \left(\frac{1}{T} \sum_i X_{i,a,p,t} \right) \\ &= \frac{1}{T} \sum_i \sum_{t=1}^{T} X_{i,a,p,t} \\ &= \frac{1}{T} s_{a,p} \end{aligned} \tag{3.8}$$

where $\bar{n}_{a,p,t}$ is the expected value of $n_{a,p,t}$ sampled randomly at time t.

Similarly,

$$N_{a,t} = \sum_i \sum_p X_{i,a,p,t} \quad \text{and} \quad S_a = \sum_i \sum_p \sum_{t=1}^{T} X_{i,a,p,t} \tag{3.9}$$

$$\begin{aligned} \Rightarrow \bar{N}_{a,t} &= \sum_{t=1}^{T} \left(\frac{1}{T} \sum_i \sum_p X_{i,a,p,t} \right) \\ &= \frac{1}{T} \sum_i \sum_p \sum_{t=1}^{T} X_{i,a,p,t} \\ &= \frac{1}{T} S_a \end{aligned} \tag{3.10}$$

where $\bar{N}_{a,t}$ is the expected value of $N_{a,t}$ sampled randomly at time t.

Dividing [3.8] by [3.10] gives

$$\frac{\bar{n}_{a,p,t}}{\bar{N}_{a,t}} = \frac{s_{a,p}}{S_a} \qquad [3.11]$$

$$\therefore \bar{n}_{a,p,t} = \bar{N}_{a,t} \frac{s_{a,p}}{S_a} . \qquad [3.12]$$

Since the distribution of the number of survivors aged a, $N_{a,t}$, is assumed to be in steady state throughout the period $0 < t \leq T$, we may estimate the expected value $\bar{N}_{a,t}$ by $N_{a,u}$, where u is the point at which the number of survivors is known ($0 < u \leq T$). Then, letting $u = T$ and $\bar{N}_{a,t} = N_{a,T}$ in [3.12] gives the required result:

$$\bar{n}_{a,p,t} \approx N_{a,T} \frac{s_{a,p}}{S_a} . \qquad [3.13]$$

3.3.3.2 Extrapolation to complete prevalence in the UK in 2008

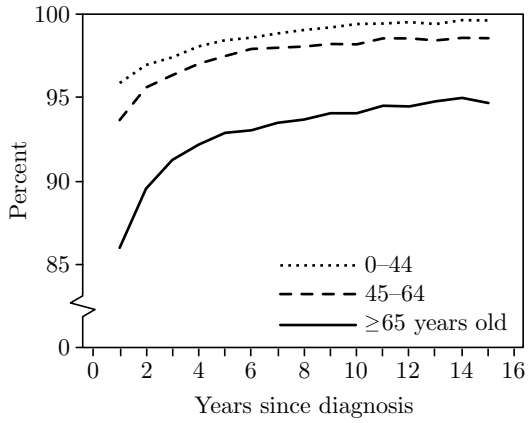
In order to use the person-time analysis pertaining to England in 2006 (as described in the previous section) to estimate the proportion of survivors in each survivorship state in the UK at the end of 2008 (cf. Chapter 2), it was assumed that the distribution of survivors between states was the same for the whole of the UK as it was for England, and did not change between 2006 and 2008.

The cancer registry dataset contained diagnoses made in the period 1990–2006. The person-time analysis for 2006 therefore completely described only those survivors diagnosed up to 16 years previously (16-year prevalence). Some extrapolation was therefore required to extend the analysis to complete prevalence. By definition, survivors more than 16 years beyond diagnosis are in either temporal phase B or D. It was assumed that the distribution of survivors between these two phases was the same for those more than 16 years beyond diagnosis as it was for those between 15 and 16 years beyond diagnosis. Figure 3.4 shows the estimated proportion of female cancer survivors in phase D, by time since diagnosis up to 16 years – it can be seen that this proportion is quite steady for time since diagnosis greater than 10 years in all age groups, and therefore the above assumption was considered reasonable.

It was also assumed that the relative numbers of survivors with an acute health service utilisation intensity of ‘high’, ‘low’, or ‘none’ in each temporal phase B or D were the same for survivors greater than 16 years beyond diagnosis as they were for those in that phase no more than 16 years beyond diagnosis.

As noted in section 2.4, the number of male lung cancer survivors at least 20 years beyond diagnosis at the end of 2008 was likely to have been overestimated (possibly by up to 8,000). For consistency, the analysis in this chapter used the same prevalence figures presented in Chapter 2, and so the number of male lung cancer survivors more than one year beyond diagnosis (phases B and D combined) is also likely to be an overestimate.

Figure 3.4. Proportion of female survivors (all malignant neoplasms) in temporal phase D, by time since diagnosis and attained age.



3.4 Results

3.4.1 Hospital activity among cancer survivors in England in 2006

Figures 3.5 and 3.6, and the corresponding Tables 3.2–3.9, display the mean number of days (per 100 person-days) spent admitted to hospital by cancer survivors in England during 2006, by attained age and time since diagnosis. Results are presented for cancer related and non-cancer related episodes of care separately, but are omitted if less than 20 person-years of prevalence contributed to the estimate.

Figure 3.5 shows that, for each age group, the number of days spent admitted to hospital for cancer related episodes of care (per 100 person-days) was highest in the first year after diagnosis. For both male and female colorectal cancer survivors, it was approximately 7 times higher than in the period 1–5 years after diagnosis, and 35 times higher than in the period ≥5 years after diagnosis. Similarly, for lung cancer survivors it was approximately 5 times higher than in the period 1–5 years after diagnosis, and 40 times higher than in the period ≥5 years after diagnosis. Although the first year after diagnosis contained by far the highest levels of cancer related hospital activity,

cancer survivors also experienced a significant amount of hospitalisation between one and five years after diagnosis.

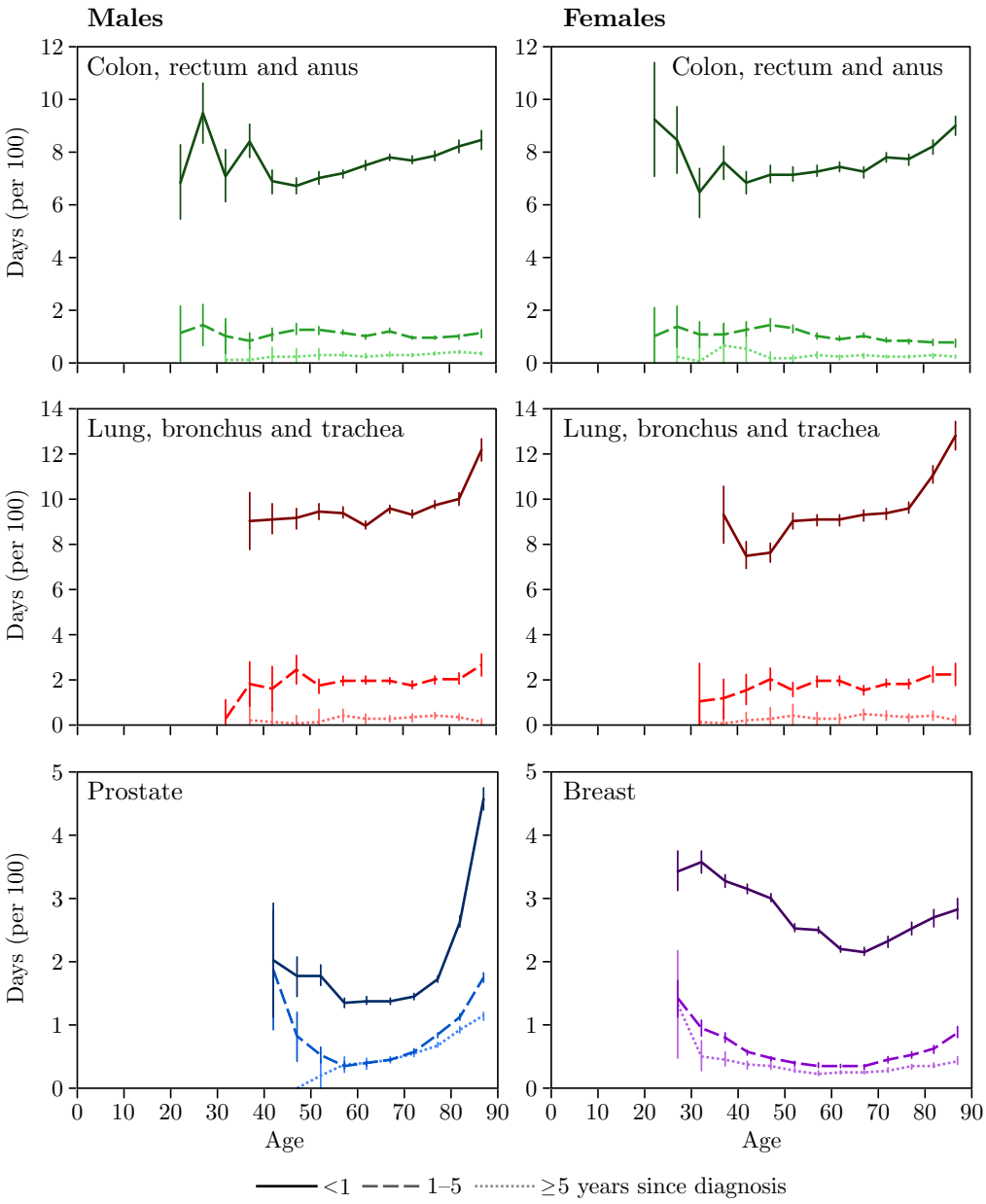
The number of days spent admitted to hospital for cancer related episodes of care (per 100 person-days) in the first year after diagnosis was significantly higher among survivors over the age of 75, compared with younger survivors. This is most noticeable for survivors of prostate cancer – those aged at least 85 years spent around three times as many days admitted to hospital for cancer related care in the first year after diagnosis, compared with those aged between 70 and 74 years.

In the first five years after diagnosis, the number of days admitted to hospital for cancer related episodes of care (per 100 person-days) was highest among lung cancer survivors; it was lowest among prostate and female breast cancer survivors. In the period more than five years after diagnosis there was little difference in the proportion of time spent admitted to hospital for cancer related care between survivors of different cancer types.

Among prostate cancer survivors, the relatively young and relatively old age groups spent a higher proportion of time admitted to hospital for cancer related episodes in the first five years after diagnosis than the middle age groups (50–75 years). A similar trend was observed among female breast cancer survivors.

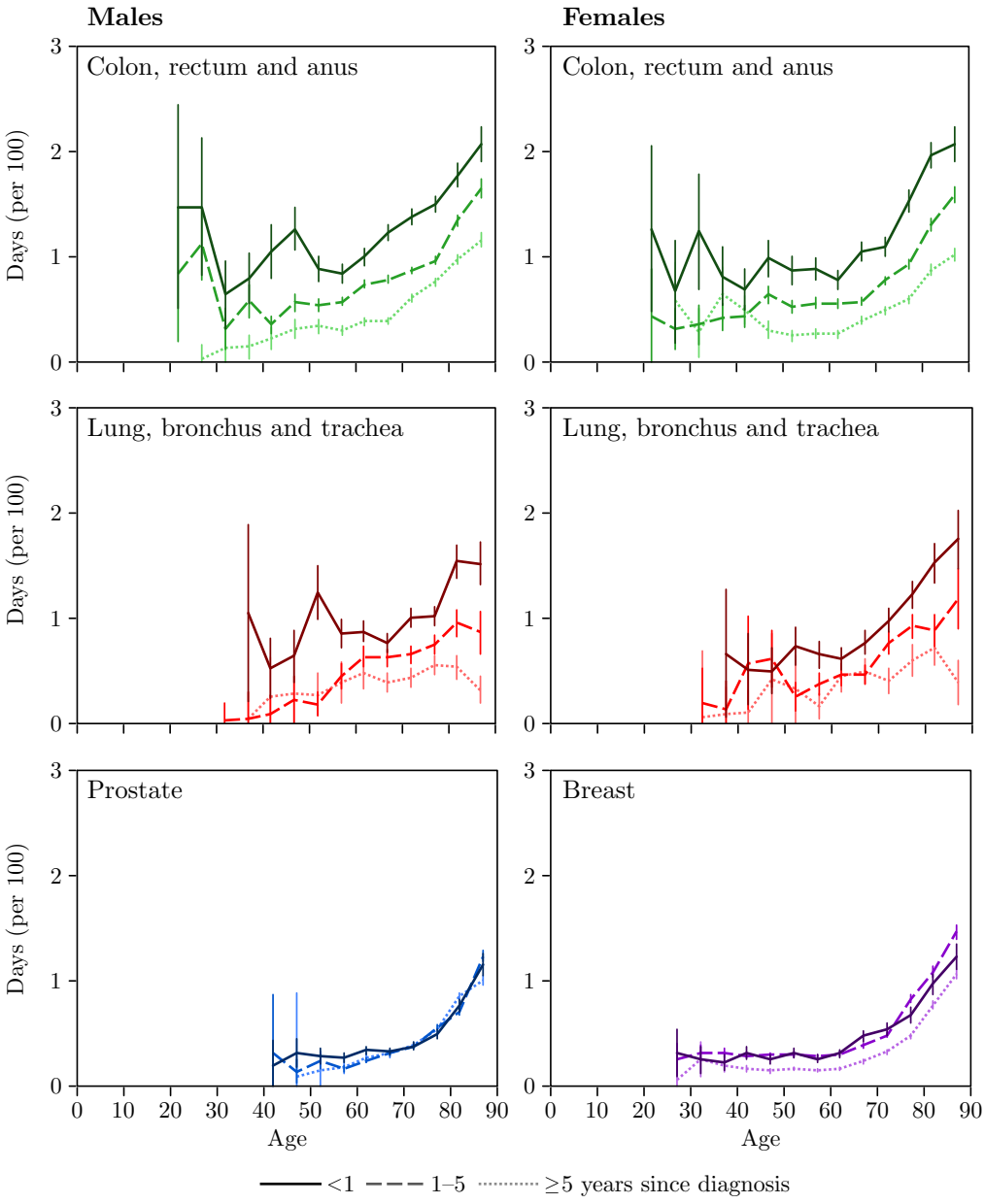
The number of days spent admitted to hospital for non-cancer related episodes of care (per 100 person-days) generally increased as attained age increased above 60 years, and showed no definitive age association among younger survivors (Figure 3.6). It generally increased slightly as time since diagnosis decreased among survivors of colorectal and lung cancer, but there was little difference in each time since diagnosis band among survivors of prostate and female breast cancers.

Figure 3.5. Mean ***cancer related*** admitted patient hospital activity among cancer survivors, England, 2006. Number of days admitted to hospital per 100 person-days (with 95% confidence intervals), by 5-year attained age group and time since diagnosis.



Results are not plotted if the total person-time of prevalence that contributed was less than 20 years.

Figure 3.6. Mean ***non-cancer related*** admitted patient hospital activity among cancer survivors, England, 2006. Number of days admitted to hospital per 100 person-days (with 95% confidence intervals), by 5-year attained age group and time since diagnosis.



Results are not plotted if the total person-time of prevalence that contributed was less than 20 years.

Table 3.2. Mean cancer related admitted patient hospital activity among survivors of colon, rectum and anus cancers, England, 2006. Number of days in hospital per 100 person-days (95% confidence interval), by attained age and time since diagnosis.

	Time since diagnosis		
Age group	<1 year	1–5 years	≥5 years
Males			
0–4			
5–9			
10–14			
15–19			
20–24	6.85 (5.44–8.25)	1.11 (0.08–2.15)	
25–29	9.48 (8.36–10.60)	1.43 (0.64–2.22)	
30–34	7.10 (6.11–8.09)	1.02 (0.34–1.70)	0.11 (0.00–1.43)
35–39	8.41 (7.78–9.04)	0.83 (0.50–1.15)	0.11 (0.00–0.58)
40–44	6.88 (6.42–7.34)	1.09 (0.84–1.34)	0.24 (0.00–0.59)
45–49	6.70 (6.39–7.00)	1.28 (1.09–1.47)	0.25 (0.00–0.51)
50–54	7.04 (6.81–7.28)	1.23 (1.09–1.37)	0.32 (0.13–0.50)
55–59	7.17 (7.00–7.34)	1.14 (1.05–1.24)	0.32 (0.21–0.42)
60–64	7.49 (7.34–7.65)	0.99 (0.91–1.06)	0.24 (0.16–0.32)
65–69	7.80 (7.65–7.94)	1.22 (1.14–1.29)	0.31 (0.24–0.38)
70–74	7.69 (7.55–7.84)	0.97 (0.90–1.04)	0.30 (0.24–0.35)
75–79	7.86 (7.70–8.03)	0.94 (0.88–1.01)	0.33 (0.27–0.38)
80–84	8.23 (7.99–8.47)	0.98 (0.89–1.07)	0.38 (0.32–0.44)
≥85	8.47 (8.11–8.83)	1.10 (0.96–1.24)	0.37 (0.30–0.44)
Females			
0–4			
5–9			
10–14			
15–19			
20–24	9.23 (7.07–11.39)	1.01 (0.00–2.06)	
25–29	8.47 (7.21–9.72)	1.36 (0.59–2.14)	0.23 (0.00–1.28)
30–34	6.45 (5.51–7.40)	1.07 (0.60–1.54)	0.04 (0.00–0.57)
35–39	7.60 (6.97–8.24)	1.09 (0.69–1.49)	0.63 (0.00–1.47)
40–44	6.84 (6.44–7.24)	1.27 (1.00–1.54)	0.51 (0.03–0.99)
45–49	7.14 (6.81–7.47)	1.45 (1.21–1.69)	0.17 (0.00–0.38)
50–54	7.15 (6.88–7.42)	1.29 (1.12–1.46)	0.16 (0.00–0.32)
55–59	7.28 (7.08–7.49)	0.99 (0.88–1.10)	0.27 (0.14–0.40)
60–64	7.43 (7.24–7.63)	0.90 (0.80–0.99)	0.21 (0.12–0.31)
65–69	7.23 (7.05–7.41)	1.02 (0.92–1.11)	0.26 (0.18–0.35)
70–74	7.79 (7.60–7.97)	0.84 (0.75–0.92)	0.24 (0.17–0.31)
75–79	7.71 (7.52–7.91)	0.83 (0.74–0.91)	0.24 (0.18–0.31)
80–84	8.19 (7.93–8.46)	0.78 (0.68–0.89)	0.26 (0.19–0.34)
≥85	9.02 (8.66–9.38)	0.74 (0.61–0.88)	0.21 (0.14–0.27)

Results are omitted if the total person-time of prevalence that contributed was less than 20 years.

Table 3.3. Mean non-cancer related admitted patient hospital activity among survivors of colon, rectum and anus cancers, England, 2006. Number of days in hospital per 100 person-days (95% confidence interval), by attained age and time since diagnosis.

	Time since diagnosis		
Age group	<1 year	1–5 years	≥5 years
Males			
0–4			
5–9			
10–14			
15–19			
20–24	1.47 (0.51–2.44)	0.84 (0.20–1.48)	
25–29	1.47 (0.82–2.13)	1.12 (0.78–1.47)	0.03 (0.00–0.17)
30–34	0.64 (0.31–0.97)	0.31 (0.11–0.51)	0.13 (0.00–0.37)
35–39	0.79 (0.55–1.04)	0.58 (0.42–0.73)	0.14 (0.03–0.26)
40–44	1.05 (0.79–1.30)	0.35 (0.27–0.43)	0.22 (0.12–0.32)
45–49	1.26 (1.06–1.46)	0.57 (0.48–0.65)	0.31 (0.23–0.40)
50–54	0.89 (0.77–1.01)	0.54 (0.48–0.59)	0.34 (0.27–0.41)
55–59	0.83 (0.75–0.92)	0.57 (0.53–0.61)	0.30 (0.26–0.34)
60–64	1.00 (0.92–1.08)	0.74 (0.70–0.77)	0.38 (0.35–0.42)
65–69	1.23 (1.15–1.31)	0.79 (0.75–0.82)	0.39 (0.36–0.42)
70–74	1.38 (1.30–1.45)	0.87 (0.84–0.90)	0.61 (0.57–0.64)
75–79	1.50 (1.42–1.58)	0.96 (0.93–1.00)	0.76 (0.72–0.79)
80–84	1.77 (1.66–1.88)	1.34 (1.29–1.39)	0.98 (0.93–1.02)
≥85	2.07 (1.90–2.23)	1.65 (1.57–1.73)	1.16 (1.10–1.23)
Females			
0–4			
5–9			
10–14			
15–19			
20–24	1.26 (0.47–2.05)	0.43 (0.00–0.89)	
25–29	0.67 (0.18–1.16)	0.31 (0.11–0.51)	0.58 (0.21–0.95)
30–34	1.24 (0.69–1.79)	0.35 (0.16–0.54)	0.28 (0.05–0.51)
35–39	0.82 (0.54–1.09)	0.42 (0.30–0.55)	0.64 (0.38–0.91)
40–44	0.69 (0.49–0.89)	0.44 (0.33–0.54)	0.49 (0.36–0.62)
45–49	0.98 (0.81–1.16)	0.64 (0.55–0.72)	0.30 (0.23–0.38)
50–54	0.87 (0.73–1.01)	0.52 (0.46–0.58)	0.25 (0.20–0.30)
55–59	0.89 (0.79–1.00)	0.56 (0.51–0.61)	0.26 (0.22–0.31)
60–64	0.78 (0.69–0.88)	0.55 (0.51–0.59)	0.26 (0.23–0.30)
65–69	1.05 (0.96–1.14)	0.57 (0.53–0.61)	0.39 (0.36–0.43)
70–74	1.10 (1.01–1.18)	0.77 (0.73–0.81)	0.49 (0.45–0.53)
75–79	1.53 (1.43–1.63)	0.93 (0.88–0.97)	0.59 (0.55–0.63)
80–84	1.96 (1.84–2.08)	1.31 (1.25–1.37)	0.87 (0.82–0.92)
≥85	2.07 (1.91–2.23)	1.59 (1.51–1.67)	1.02 (0.96–1.07)

Results are omitted if the total person-time of prevalence that contributed was less than 20 years.

Table 3.4. Mean cancer related admitted patient hospital activity among survivors of lung, bronchus and trachea cancers, England, 2006. Number of days in hospital per 100 person-days (95% confidence interval), by attained age and time since diagnosis.

	Time since diagnosis		
Age group	<1 year	1–5 years	≥5 years
Males			
0–4			
5–9			
10–14			
15–19			
20–24			
25–29			
30–34		0.25 (0.00–1.08)	
35–39	9.02 (7.76–10.27)	1.81 (0.86–2.77)	0.21 (0.00–1.59)
40–44	9.12 (8.45–9.79)	1.60 (0.64–2.57)	0.10 (0.00–0.74)
45–49	9.15 (8.69–9.61)	2.43 (1.78–3.08)	0.07 (0.00–0.44)
50–54	9.46 (9.13–9.80)	1.72 (1.42–2.02)	0.16 (0.00–0.71)
55–59	9.40 (9.18–9.63)	1.93 (1.72–2.15)	0.43 (0.15–0.72)
60–64	8.83 (8.65–9.01)	1.98 (1.81–2.15)	0.29 (0.09–0.50)
65–69	9.58 (9.41–9.76)	1.95 (1.79–2.12)	0.29 (0.15–0.44)
70–74	9.32 (9.14–9.49)	1.77 (1.61–1.94)	0.31 (0.16–0.45)
75–79	9.76 (9.56–9.95)	1.99 (1.82–2.17)	0.39 (0.26–0.53)
80–84	10.04 (9.76–10.32)	2.05 (1.80–2.30)	0.35 (0.20–0.51)
≥85	12.20 (11.72–12.68)	2.65 (2.19–3.11)	0.12 (0.00–0.25)
Females			
0–4			
5–9			
10–14			
15–19			
20–24			
25–29			
30–34		1.01 (0.00–2.71)	0.11 (0.00–1.03)
35–39	9.30 (8.05–10.54)	1.15 (0.26–2.04)	0.04 (0.00–0.76)
40–44	7.50 (6.90–8.10)	1.56 (0.92–2.20)	0.17 (0.00–0.54)
45–49	7.63 (7.24–8.02)	2.04 (1.56–2.52)	0.26 (0.00–0.76)
50–54	9.03 (8.69–9.36)	1.54 (1.22–1.86)	0.38 (0.00–0.88)
55–59	9.07 (8.83–9.31)	1.92 (1.69–2.15)	0.26 (0.01–0.52)
60–64	9.07 (8.85–9.30)	1.96 (1.75–2.17)	0.28 (0.06–0.51)
65–69	9.28 (9.06–9.49)	1.53 (1.34–1.71)	0.45 (0.22–0.67)
70–74	9.35 (9.13–9.58)	1.84 (1.64–2.04)	0.40 (0.18–0.63)
75–79	9.60 (9.34–9.85)	1.81 (1.59–2.03)	0.33 (0.16–0.51)
80–84	11.08 (10.71–11.45)	2.23 (1.90–2.55)	0.43 (0.20–0.65)
≥85	12.81 (12.19–13.42)	2.23 (1.77–2.69)	0.17 (0.00–0.40)

Results are omitted if the total person-time of prevalence that contributed was less than 20 years.

Table 3.5. Mean non-cancer related admitted patient hospital activity among survivors of lung, bronchus and trachea cancers, England, 2006. Number of days in hospital per 100 person-days (95% confidence interval), by attained age and time since diagnosis.

	Time since diagnosis		
Age group	<1 year	1–5 years	≥5 years
Males			
0–4			
5–9			
10–14			
15–19			
20–24			
25–29			
30–34		0.03 (0.00–0.20)	
35–39	1.05 (0.21–1.89)	0.04 (0.00–0.30)	0.04 (0.00–0.21)
40–44	0.52 (0.23–0.80)	0.08 (0.00–0.26)	0.25 (0.00–0.63)
45–49	0.64 (0.38–0.89)	0.22 (0.00–0.45)	0.28 (0.00–0.69)
50–54	1.24 (0.99–1.50)	0.19 (0.07–0.30)	0.27 (0.07–0.48)
55–59	0.85 (0.72–0.99)	0.44 (0.32–0.56)	0.39 (0.19–0.59)
60–64	0.88 (0.77–0.98)	0.63 (0.54–0.73)	0.47 (0.33–0.62)
65–69	0.77 (0.68–0.86)	0.62 (0.54–0.71)	0.38 (0.29–0.47)
70–74	1.00 (0.91–1.09)	0.66 (0.57–0.74)	0.44 (0.35–0.52)
75–79	1.02 (0.93–1.11)	0.75 (0.66–0.84)	0.55 (0.46–0.64)
80–84	1.54 (1.38–1.69)	0.95 (0.82–1.08)	0.53 (0.42–0.64)
≥85	1.52 (1.32–1.72)	0.86 (0.65–1.07)	0.32 (0.20–0.44)
Females			
0–4			
5–9			
10–14			
15–19			
20–24			
25–29			
30–34		0.20 (0.00–0.52)	0.06 (0.00–0.68)
35–39	0.66 (0.05–1.27)	0.14 (0.00–0.41)	0.09 (0.00–0.34)
40–44	0.51 (0.18–0.85)	0.57 (0.13–1.02)	0.10 (0.00–0.32)
45–49	0.50 (0.28–0.72)	0.61 (0.34–0.89)	0.41 (0.00–0.85)
50–54	0.73 (0.55–0.92)	0.25 (0.11–0.38)	0.32 (0.00–0.65)
55–59	0.66 (0.53–0.79)	0.38 (0.27–0.48)	0.16 (0.05–0.28)
60–64	0.61 (0.50–0.72)	0.47 (0.36–0.57)	0.45 (0.30–0.59)
65–69	0.77 (0.66–0.88)	0.46 (0.37–0.56)	0.49 (0.37–0.62)
70–74	0.98 (0.86–1.09)	0.76 (0.65–0.87)	0.40 (0.29–0.52)
75–79	1.22 (1.09–1.35)	0.92 (0.81–1.04)	0.60 (0.45–0.74)
80–84	1.52 (1.34–1.71)	0.88 (0.74–1.03)	0.72 (0.56–0.89)
≥85	1.75 (1.47–2.03)	1.19 (0.90–1.47)	0.39 (0.17–0.60)

Results are omitted if the total person-time of prevalence that contributed was less than 20 years.

Table 3.6. Mean cancer related admitted patient hospital activity among survivors of prostate cancer, England, 2006. Number of days in hospital per 100 person-days (95% confidence interval), by attained age and time since diagnosis.

	Time since diagnosis		
Age group	<1 year	1–5 years	≥5 years
Males			
0–4			
5–9			
10–14			
15–19			
20–24			
25–29			
30–34			
35–39			
40–44	2.02 (1.13–2.92)	1.87 (0.92–2.81)	
45–49	1.77 (1.46–2.08)	0.81 (0.43–1.20)	0.00 (0.00–0.03*)
50–54	1.78 (1.62–1.94)	0.52 (0.39–0.65)	0.20 (0.00–0.44)
55–59	1.34 (1.27–1.42)	0.36 (0.30–0.41)	0.37 (0.24–0.50)
60–64	1.38 (1.32–1.44)	0.39 (0.35–0.43)	0.39 (0.31–0.46)
65–69	1.38 (1.33–1.43)	0.45 (0.42–0.48)	0.46 (0.41–0.51)
70–74	1.44 (1.39–1.49)	0.58 (0.55–0.61)	0.55 (0.51–0.59)
75–79	1.72 (1.66–1.78)	0.84 (0.80–0.88)	0.68 (0.64–0.72)
80–84	2.63 (2.54–2.73)	1.13 (1.08–1.18)	0.93 (0.88–0.97)
≥85	4.57 (4.41–4.74)	1.75 (1.67–1.83)	1.15 (1.08–1.21)

Results are omitted if the total person-time of prevalence that contributed was less than 20 years.
*Wilson (score) confidence interval.

Table 3.7. Mean non-cancer related admitted patient hospital activity among survivors of prostate cancer, England, 2006. Number of days in hospital per 100 person-days (95% confidence interval), by attained age and time since diagnosis.

	Time since diagnosis		
Age group	<1 year	1–5 years	≥5 years
Males			
0–4			
5–9			
10–14			
15–19			
20–24			
25–29			
30–34			
35–39			
40–44	0.19 (0.00–0.44)	0.31 (0.00–0.86)	
45–49	0.31 (0.18–0.44)	0.14 (0.02–0.25)	0.09 (0.00–0.89)
50–54	0.28 (0.21–0.36)	0.23 (0.17–0.30)	0.15 (0.00–0.33)
55–59	0.26 (0.22–0.31)	0.16 (0.13–0.19)	0.18 (0.11–0.26)
60–64	0.34 (0.30–0.38)	0.24 (0.22–0.26)	0.27 (0.23–0.32)
65–69	0.33 (0.30–0.36)	0.31 (0.29–0.33)	0.31 (0.27–0.34)
70–74	0.38 (0.34–0.41)	0.38 (0.36–0.40)	0.39 (0.37–0.42)
75–79	0.49 (0.45–0.53)	0.56 (0.53–0.58)	0.54 (0.52–0.57)
80–84	0.76 (0.70–0.83)	0.71 (0.68–0.74)	0.85 (0.81–0.88)
≥85	1.16 (1.05–1.26)	1.23 (1.18–1.29)	1.00 (0.96–1.04)

Results are omitted if the total person-time of prevalence that contributed was less than 20 years.

Table 3.8. Mean cancer related admitted patient hospital activity among female survivors of breast cancer, England, 2006. Number of days in hospital per 100 person-days (95% confidence interval), by attained age and time since diagnosis.

	Time since diagnosis		
Age group	<1 year	1–5 years	≥5 years
Females			
0–4			
5–9			
10–14			
15–19			
20–24			
25–29	3.43 (3.12–3.74)	1.41 (1.12–1.70)	1.33 (0.48–2.18)
30–34	3.57 (3.40–3.74)	0.95 (0.82–1.08)	0.51 (0.27–0.74)
35–39	3.27 (3.18–3.37)	0.79 (0.72–0.87)	0.46 (0.35–0.56)
40–44	3.14 (3.07–3.21)	0.56 (0.51–0.61)	0.36 (0.30–0.42)
45–49	3.01 (2.95–3.07)	0.47 (0.43–0.51)	0.34 (0.30–0.39)
50–54	2.53 (2.48–2.59)	0.38 (0.35–0.42)	0.28 (0.24–0.31)
55–59	2.49 (2.44–2.55)	0.35 (0.32–0.38)	0.23 (0.21–0.26)
60–64	2.21 (2.15–2.26)	0.34 (0.31–0.37)	0.25 (0.22–0.27)
65–69	2.16 (2.10–2.22)	0.34 (0.30–0.38)	0.25 (0.22–0.28)
70–74	2.31 (2.23–2.40)	0.45 (0.41–0.50)	0.28 (0.25–0.31)
75–79	2.52 (2.41–2.62)	0.53 (0.47–0.58)	0.34 (0.30–0.38)
80–84	2.69 (2.55–2.82)	0.61 (0.54–0.68)	0.36 (0.31–0.40)
≥85	2.83 (2.67–2.99)	0.88 (0.79–0.96)	0.43 (0.37–0.49)

Results are omitted if the total person-time of prevalence that contributed was less than 20 years.

Table 3.9. Mean non-cancer related admitted patient hospital activity among female survivors of breast cancer, England, 2006. Number of days in hospital per 100 person-days (95% confidence interval), by attained age and time since diagnosis.

	Time since diagnosis		
Age group	<1 year	1–5 years	≥5 years
Females			
0–4			
5–9			
10–14			
15–19			
20–24			
25–29	0.31 (0.08–0.54)	0.25 (0.10–0.41)	0.06 (0.00–0.26)
30–34	0.25 (0.12–0.37)	0.32 (0.24–0.39)	0.25 (0.09–0.42)
35–39	0.22 (0.15–0.30)	0.31 (0.27–0.35)	0.19 (0.13–0.26)
40–44	0.31 (0.26–0.37)	0.29 (0.26–0.32)	0.16 (0.13–0.20)
45–49	0.25 (0.20–0.29)	0.29 (0.27–0.32)	0.15 (0.12–0.17)
50–54	0.32 (0.27–0.36)	0.29 (0.27–0.31)	0.16 (0.14–0.18)
55–59	0.25 (0.22–0.29)	0.28 (0.26–0.30)	0.15 (0.13–0.16)
60–64	0.31 (0.27–0.35)	0.30 (0.28–0.32)	0.17 (0.15–0.18)
65–69	0.47 (0.42–0.52)	0.38 (0.36–0.41)	0.23 (0.22–0.25)
70–74	0.54 (0.48–0.60)	0.48 (0.46–0.51)	0.32 (0.30–0.34)
75–79	0.67 (0.60–0.74)	0.83 (0.79–0.87)	0.47 (0.45–0.50)
80–84	0.97 (0.87–1.07)	1.08 (1.04–1.13)	0.77 (0.73–0.80)
≥85	1.23 (1.11–1.35)	1.47 (1.40–1.53)	1.06 (1.02–1.11)

Results are omitted if the total person-time of prevalence that contributed was less than 20 years.

3.4.2 Survivorship states in the UK in 2008

Tables 3.10–3.14 present estimates of the number of UK cancer survivors in different survivorship states at the end of 2008. Each table is for a different cancer type. Tables 3.15–3.19 present the same data as percentages. To aid interpretation of the results in these tables, Figures 3.9–3.13 are also included. These are bespoke graphics which represent cancer survivors by a set of square tiles, each with four horizontal divisions marking the proportion of survivors in each of the four temporal phases (defined according to time since diagnosis and time until death – see Figure 3.3), and three vertical divisions marking the proportion of survivors in each category of acute health service utilisation (none, low or high intensity – see section 3.3.3.1). In essence, these figures are a variation on the standard histogram chart and employ the same principle that the area of each block is proportional to the quantity represented. Each square tile represents the survivors of a given cancer type, sex and broad age group, and the density of overlaid male or female person icons is approximately proportional to the total number of survivors represented by the tile – each person icon represents 1,000 cancer survivors.

The design of Figures 3.9–3.13 was inspired by the work of the Dutch artist, Piet Mondrian (1872–1944). Mondrian was a founding contributor to the art review *De Stijl* ("The Style") which was first published in the Netherlands in October 1917 and from which grew the neo-plasticism art movement (Elgar, 1968). Manifestos printed in early editions of *De Stijl* set out the principles of neo-plasticism:

> "...'painting must be made to submit to the horizontal-vertical order, which excludes the diagonal and the curve; colours must be limited to the three primary colours and the three non-colours white, black and grey, which must never be either mixed or superimposed (cookery)...What we want is a new aesthetic based on pure relationships of lines and pure colours, because only pure relationships between pure constructive elements can result in pure beauty.'"
>
> (Elgar, 1968: p.93)

An example of a composition by Mondrian in the neo-plastic style is shown in Figure 3.7.

Mondrian's ideas and philosophy have been adapted for this results section to provide a visually attractive way of displaying a large amount of data in a clean and simple way, allowing for an immediate high level analysis, as well as facilitating in-depth study and easy comparison across cancer types, age groups and sexes. Figure 3.8 provides a

reference key for Figures 3.9–3.13, and indicates which temporal phase and health service utilisation categories correspond to each section of the tile. It should be noted that, in Figures 3.9–3.13, the proportion of survivors in temporal phase D and with a 'high' level of acute health service utilisation (i.e. the rightmost area of the bottom row) may be too small to be visible. Similarly, the proportion of breast cancer survivors in temporal phase A (i.e. the top row) may also be too small to be visible.

Figure 3.7. Composition with Red, Yellow, Blue and Black by Piet Mondrian (1921). Oil on canvas 59.5 × 59.5 cm. Gemeentemuseum, The Hague, Netherlands.

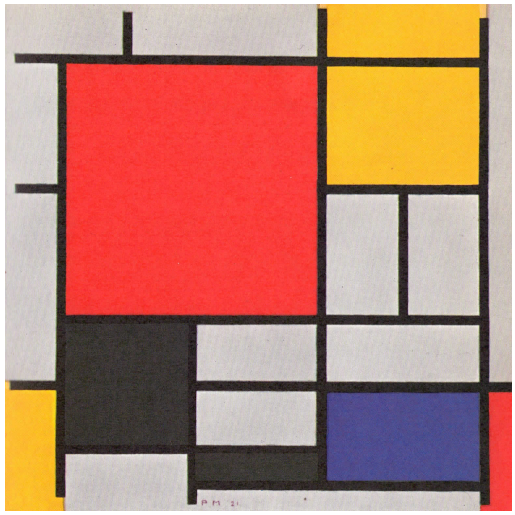


Figure 3.8. Neo-plastic reference key for Figures 3.9–3.13.

Temporal phase	Health service utilisation intensity: None	Low	High
A	A: None	A: Low	A: High
B	B: None	B: Low	B: High
C	C: None	C: Low	C: High
D	D: None	D: Low	D: High

At the end of 2008, approximately 81% of all male, and 87% of all female, cancer survivors were more than one year beyond diagnosis and more than one year from

death (Table 3.19). There was little variation in these proportions by attained age, except for male survivors aged under 45 years for whom this proportion rose to 90%. The proportion of lung cancer survivors who were more than one year beyond diagnosis and more than one year from death was markedly lower (72% of male survivors and 64% of female survivors), as shown in Table 3.16. There was little acute health service utilisation among cancer survivors in this temporal phase, regardless of attained age or type of cancer. Overall, 1.57 million of the 2 million cancer survivors in the UK at the end of 2008 (78%) were more than one year beyond diagnosis and more than one year from death, and had no cancer related hospital admissions in this temporal phase.

Of the estimated two million cancer survivors in the UK (Chapter 2), 240,000 were in the 'low' acute health service utilisation category (14.8% of male survivors and 10.0% of female survivors) – see Tables 3.14 and 3.19. 61,000 (4.0% of male survivors and 2.4% of female survivors) were in the 'high' category. The proportion of lung cancer survivors in the 'high' category was much larger (7.2% of males and 9.3% of females), particularly in the age group 45–64 years where the proportion rose to around 12% (Table 3.16). Conversely, only 1.2% of female breast cancer survivors had a 'high' level of acute health service utilisation (Table 3.18).

Table 3.14 also shows that 147,000 cancer survivors were in the last year of their life (9.0% of male survivors and 6.2% of female survivors). It was these survivors who had the highest intensity of cancer related acute health service utilisation, particularly those who were also in the first year after diagnosis. 41% of the 41,000 cancer survivors who were less than one year beyond diagnosis and less than one year from death had a high intensity of acute health service utilisation, compared with 19% of the 106,000 survivors who were in the last year of their life but more than one year beyond diagnosis.

Table 3.10. Number of colon, rectum and anus cancer survivors in the UK, 2008. By sex, attained age, temporal phase of survivorship* and intensity of cancer related health service utilisation†.

	Age group															
	0–44				45–64				≥65				All ages			
	Intensity of health service use				Intensity of health service use				Intensity of health service use				Intensity of health service use			
Temporal phase	None	Low	High	Total	None	Low	High	Total	None	Low	High	Total	None	Low	High	Total
Males																
A	<10	20	30	60	80	260	270	610	530	710	840	2,080	610	990	1,150	2,750
B	10	40	30	80	290	510	270	1,060	4,120	2,130	1,070	7,320	4,420	2,680	1,360	8,460
C	130	150	80	360	1,050	1,820	790	3,670	2,590	3,650	1,590	7,840	3,770	5,620	2,470	11,860
D	1,490	90	<10	1,590	18,760	1,430	160	20,350	72,400	4,700	410	77,520	92,650	6,220	580	99,450
Total	1,640	310	150	2,090	20,180	4,010	1,500	25,690	79,640	11,200	3,910	94,750	101,460	15,520	5,560	122,530
Females																
A	<10	10	30	50	50	140	190	390	510	500	750	1,750	570	650	970	2,180
B	20	30	30	80	180	300	200	680	4,010	1,470	820	6,300	4,210	1,800	1,060	7,060
C	110	150	70	330	710	1,260	540	2,510	2,310	2,790	1,190	6,290	3,130	4,200	1,800	9,120
D	1,550	110	10	1,680	15,140	910	100	16,150	73,780	3,010	310	77,090	90,470	4,030	420	94,910
Total	1,690	300	150	2,130	16,080	2,610	1,030	19,720	80,600	7,760	3,070	91,430	98,370	10,670	4,250	113,290

*†As defined in section 3.3.3.

Table 3.11. Number of lung, bronchus and trachea cancer survivors in the UK, 2008. By sex, attained age, temporal phase of survivorship* and intensity of cancer related health service utilisation†.

	Age group															
	0–44				45–64				≥65				All ages			
	Intensity of health service use				Intensity of health service use				Intensity of health service use				Intensity of health service use			
Temporal phase	None	Low	High	Total	None	Low	High	Total	None	Low	High	Total	None	Low	High	Total
Males																
A	<10	30	30	70	210	660	520	1,400	830	1,420	1,260	3,500	1,050	2,120	1,810	4,980
B	<10	<10	<10	20	180	260	120	550	930	770	370	2,070	1,110	1,040	490	2,640
C	20	40	<10	70	320	630	130	1,090	830	1,090	220	2,130	1,180	1,750	360	3,290
D	280	10	<10	290	3,310	270	20	3,600	22,360	1,500	90	23,950	25,950	1,780	110	27,840
Total	310	90	50	440	4,030	1,820	790	6,640	24,950	4,780	1,940	31,660	29,280	6,690	2,770	38,740
Females																
A	<10	30	20	60	160	480	440	1,080	750	1,060	1,030	2,840	920	1,580	1,490	3,980
B	<10	10	<10	20	150	220	120	500	760	520	290	1,570	920	760	420	2,090
C	30	30	<10	70	300	570	130	1,000	780	860	200	1,850	1,110	1,460	350	2,920
D	340	30	<10	370	3,040	260	20	3,320	11,320	740	40	12,090	14,700	1,020	60	15,780
Total	380	110	40	530	3,660	1,530	710	5,900	13,600	3,180	1,560	18,350	17,640	4,820	2,320	24,780

*†As defined in section 3.3.3.

Table 3.12. Number of prostate cancer survivors in the UK, 2008. By attained age, temporal phase of survivorship* and intensity of cancer related health service utilisation†.

	Age group															
	0–44				45–64				≥65				All ages			
	Intensity of health service use				Intensity of health service use				Intensity of health service use				Intensity of health service use			
Temporal phase	None	Low	High	Total	None	Low	High	Total	None	Low	High	Total	None	Low	High	Total
A	<10	<10	<10	<10	80	130	70	280	1,110	890	740	2,740	1,190	1,020	820	3,030
B	<10	<10	<10	<10	220	350	180	750	7,800	5,860	3,210	16,880	8,020	6,210	3,390	17,630
C	20	10	<10	40	5,020	3,590	140	8,740	17,190	8,370	600	26,160	22,230	11,970	740	34,940
D	120	10	<10	130	22,830	1,830	80	24,740	152,850	18,960	1,150	172,970	175,810	20,800	1,230	197,840
Total	140	30	10	180	28,150	5,890	470	34,510	178,960	34,080	5,700	218,740	207,250	40,000	6,180	253,440

*†As defined in section 3.3.3.

Table 3.13. Number of female breast cancer survivors in the UK, 2008. By attained age, temporal phase of survivorship* and intensity of cancer related health service utilisation†.

	Age group															
	0–44				45–64				≥65				All ages			
	Intensity of health service use				Intensity of health service use				Intensity of health service use				Intensity of health service use			
Temporal phase	None	Low	High	Total	None	Low	High	Total	None	Low	High	Total	None	Low	High	Total
A	30	60	30	120	130	250	140	520	830	480	360	1,670	990	790	530	2,310
B	150	370	190	710	1,260	2,110	1,020	4,390	11,200	4,390	2,400	17,990	12,610	6,870	3,620	23,100
C	1,430	3,060	210	4,700	7,610	11,430	610	19,660	8,250	7,020	500	15,770	17,290	21,520	1,320	40,120
D	17,950	1,870	80	19,890	174,890	8,240	380	183,510	268,030	11,200	840	280,070	460,870	21,300	1,300	483,470
Total	19,560	5,360	510	25,430	183,890	22,030	2,150	208,080	288,300	23,090	4,100	315,490	491,750	50,480	6,760	549,000

*†As defined in section 3.3.3.

Table 3.14. Number of survivors of all malignant neoplasms combined in the UK, 2008. By sex, attained age, temporal phase of survivorship* and intensity of cancer related health service utilisation†.

	Age group															
	0–44				45–64				≥65				All ages			
	Intensity of health service use				Intensity of health service use				Intensity of health service use				Intensity of health service use			
Temporal phase	None	Low	High	Total	None	Low	High	Total	None	Low	High	Total	None	Low	High	Total
Males																
A	80	210	360	650	880	2,260	2,500	5,640	4,180	6,170	6,680	17,030	5,140	8,630	9,540	23,310
B	220	360	300	880	1,960	2,680	1,790	6,430	21,210	14,480	8,060	43,750	23,390	17,520	10,140	51,060
C	2,400	2,410	1,150	5,950	10,330	10,810	2,740	23,880	24,500	19,820	4,270	48,600	37,230	33,040	8,160	78,430
D	59,200	4,020	560	63,780	142,560	12,260	1,150	155,970	398,000	45,800	2,840	446,640	599,760	62,080	4,560	666,390
Total	61,890	6,990	2,370	71,250	155,730	28,010	8,180	191,920	447,890	86,270	21,850	556,020	665,520	121,270	32,400	819,190
Females																
A	100	230	360	690	670	1,650	1,970	4,290	3,570	4,100	5,010	12,680	4,340	5,980	7,340	17,660
B	310	680	510	1,490	2,690	4,230	2,550	9,470	25,610	11,570	6,890	44,070	28,610	16,470	9,950	55,030
C	4,390	5,620	1,140	11,160	13,460	18,490	2,560	34,510	17,200	17,270	3,510	37,970	35,050	41,380	7,210	83,640
D	76,500	5,240	540	82,280	318,620	17,370	1,200	337,190	573,070	32,000	2,460	607,520	968,180	54,620	4,190	1,026,990
Total	81,300	11,770	2,550	95,620	335,440	41,730	8,280	385,460	619,440	64,940	17,870	702,250	1,036,180	118,440	28,700	1,183,330

*†As defined in section 3.3.3.

Table 3.15. Percentage of colon, rectum and anus cancer survivors in the UK, 2008. By sex, attained age, temporal phase of survivorship* and intensity of cancer related health service utilisation†.

	Age group															
	0–44				45–64				≥65				All ages			
	Intensity of health service use				Intensity of health service use				Intensity of health service use				Intensity of health service use			
Temporal phase	None	Low	High	Total	None	Low	High	Total	None	Low	High	Total	None	Low	High	Total
Males																
A	0.4	1.2	1.5	3.1	0.3	1.0	1.1	2.4	0.6	0.8	0.9	2.2	0.5	0.8	0.9	2.2
B	0.7	1.9	1.2	3.8	1.1	2.0	1.1	4.1	4.3	2.2	1.1	7.7	3.6	2.2	1.1	6.9
C	6.2	7.3	3.8	17.3	4.1	7.1	3.1	14.3	2.7	3.9	1.7	8.3	3.1	4.6	2.0	9.7
D	71.1	4.3	0.5	75.8	73.0	5.5	0.6	79.2	76.4	5.0	0.4	81.8	75.6	5.1	0.5	81.2
Total	78.3	14.7	7.0	(100.0)	78.6	15.6	5.8	(100.0)	84.0	11.8	4.1	(100.0)	82.8	12.7	4.5	(100.0)
Females																
A	0.2	0.6	1.5	2.4	0.3	0.7	1.0	2.0	0.6	0.5	0.8	1.9	0.5	0.6	0.9	1.9
B	1.0	1.3	1.4	3.7	0.9	1.5	1.0	3.5	4.4	1.6	0.9	6.9	3.7	1.6	0.9	6.2
C	5.2	7.0	3.4	15.5	3.6	6.4	2.7	12.7	2.5	3.0	1.3	6.9	2.8	3.7	1.6	8.1
D	72.6	5.2	0.6	78.5	76.8	4.6	0.5	81.9	80.7	3.3	0.3	84.3	79.9	3.6	0.4	83.8
Total	79.0	14.1	6.9	(100.0)	81.5	13.2	5.2	(100.0)	88.2	8.5	3.4	(100.0)	86.8	9.4	3.8	(100.0)

*†As defined in section 3.3.3.

Table 3.16. Percentage of lung, bronchus and trachea cancer survivors in the UK, 2008. By sex, attained age, temporal phase of survivorship* and intensity of cancer related health service utilisation†.

	Age group															
	0–44				45–64				≥65				All ages			
	Intensity of health service use				Intensity of health service use				Intensity of health service use				Intensity of health service use			
Temporal phase	None	Low	High	Total	None	Low	High	Total	None	Low	High	Total	None	Low	High	Total
Males																
A	1.7	7.6	7.0	16.3	3.2	10.0	7.9	21.1	2.6	4.5	4.0	11.1	2.7	5.5	4.7	12.8
B	0.7	1.4	1.5	3.6	2.7	3.9	1.8	8.4	2.9	2.4	1.2	6.5	2.9	2.7	1.3	6.8
C	5.0	7.9	2.1	15.0	4.9	9.5	2.0	16.4	2.6	3.4	0.7	6.7	3.0	4.5	0.9	8.5
D	62.4	2.6	0.2	65.1	49.9	4.0	0.2	54.2	70.6	4.7	0.3	75.7	67.0	4.6	0.3	71.9
Total	69.8	19.5	10.7	(100.0)	60.7	27.5	11.9	(100.0)	78.8	15.1	6.1	(100.0)	75.6	17.3	7.2	(100.0)
Females																
A	1.8	6.1	4.3	12.2	2.8	8.2	7.4	18.4	4.1	5.8	5.6	15.5	3.7	6.4	6.0	16.1
B	1.1	2.7	0.7	4.5	2.6	3.7	2.1	8.4	4.1	2.9	1.6	8.6	3.7	3.1	1.7	8.4
C	5.1	6.4	1.6	13.1	5.2	9.6	2.3	17.0	4.2	4.7	1.1	10.1	4.5	5.9	1.4	11.8
D	64.7	5.2	0.3	70.2	51.4	4.4	0.3	56.2	61.7	4.0	0.2	65.9	59.3	4.1	0.3	63.7
Total	72.6	20.4	7.0	(100.0)	62.0	25.9	12.1	(100.0)	74.1	17.3	8.5	(100.0)	71.2	19.5	9.3	(100.0)

*†As defined in section 3.3.3.

Table 3.17. Percentage of prostate cancer survivors in the UK, 2008. By attained age, temporal phase of survivorship* and intensity of cancer related health service utilisation†.

	Age group															
	0–44				45–64				≥65				All ages			
	Intensity of health service use				Intensity of health service use				Intensity of health service use				Intensity of health service use			
Temporal phase	None	Low	High	Total	None	Low	High	Total	None	Low	High	Total	None	Low	High	Total
A	0.0	0.5	1.0	1.6	0.2	0.4	0.2	0.8	0.5	0.4	0.3	1.3	0.5	0.4	0.3	1.2
B	0.8	0.5	1.3	2.6	0.6	1.0	0.5	2.2	3.6	2.7	1.5	7.7	3.2	2.5	1.3	7.0
C	13.2	7.3	2.4	22.9	14.5	10.4	0.4	25.3	7.9	3.8	0.3	12.0	8.8	4.7	0.3	13.8
D	64.5	6.8	1.7	73.0	66.2	5.3	0.2	71.7	69.9	8.7	0.5	79.1	69.4	8.2	0.5	78.1
Total	78.5	15.2	6.4	(100.0)	81.6	17.1	1.4	(100.0)	81.8	15.6	2.6	(100.0)	81.8	15.8	2.4	(100.0)

*†As defined in section 3.3.3.

Table 3.18. Percentage of female breast cancer survivors in the UK, 2008. By attained age, temporal phase of survivorship* and intensity of cancer related health service utilisation†.

	Age group															
	0–44				45–64				≥65				All ages			
	Intensity of health service use				Intensity of health service use				Intensity of health service use				Intensity of health service use			
Temporal phase	None	Low	High	Total	None	Low	High	Total	None	Low	High	Total	None	Low	High	Total
A	0.1	0.2	0.1	0.5	0.1	0.1	0.1	0.2	0.3	0.2	0.1	0.5	0.2	0.1	0.1	0.4
B	0.6	1.5	0.8	2.8	0.6	1.0	0.5	2.1	3.5	1.4	0.8	5.7	2.3	1.3	0.7	4.2
C	5.6	12.0	0.8	18.5	3.7	5.5	0.3	9.4	2.6	2.2	0.2	5.0	3.1	3.9	0.2	7.3
D	70.6	7.3	0.3	78.2	84.1	4.0	0.2	88.2	85.0	3.5	0.3	88.8	83.9	3.9	0.2	88.1
Total	76.9	21.1	2.0	(100.0)	88.4	10.6	1.0	(100.0)	91.4	7.3	1.3	(100.0)	89.6	9.2	1.2	(100.0)

*†As defined in section 3.3.3.

Table 3.19. Percentage of survivors of all malignant neoplasms combined in the UK, 2008. By sex, attained age, temporal phase of survivorship* and intensity of cancer related health service utilisation†.

	Age group															
	0–44				45–64				≥65				All ages			
	Intensity of health service use				Intensity of health service use				Intensity of health service use				Intensity of health service use			
Temporal phase	None	Low	High	Total	None	Low	High	Total	None	Low	High	Total	None	Low	High	Total
Males																
A	0.1	0.3	0.5	0.9	0.5	1.2	1.3	2.9	0.8	1.1	1.2	3.1	0.6	1.1	1.2	2.8
B	0.3	0.5	0.4	1.2	1.0	1.4	0.9	3.4	3.8	2.6	1.4	7.9	2.9	2.1	1.2	6.2
C	3.4	3.4	1.6	8.3	5.4	5.6	1.4	12.4	4.4	3.6	0.8	8.7	4.5	4.0	1.0	9.6
D	83.1	5.6	0.8	89.5	74.3	6.4	0.6	81.3	71.6	8.2	0.5	80.3	73.2	7.6	0.6	81.3
Total	86.9	9.8	3.3	(100.0)	81.1	14.6	4.3	(100.0)	80.6	15.5	3.9	(100.0)	81.2	14.8	4.0	(100.0)
Females																
A	0.1	0.2	0.4	0.7	0.2	0.4	0.5	1.1	0.5	0.6	0.7	1.8	0.4	0.5	0.6	1.5
B	0.3	0.7	0.5	1.6	0.7	1.1	0.7	2.5	3.6	1.6	1.0	6.3	2.4	1.4	0.8	4.7
C	4.6	5.9	1.2	11.7	3.5	4.8	0.7	9.0	2.4	2.5	0.5	5.4	3.0	3.5	0.6	7.1
D	80.0	5.5	0.6	86.0	82.7	4.5	0.3	87.5	81.6	4.6	0.3	86.5	81.8	4.6	0.4	86.8
Total	85.0	12.3	2.7	(100.0)	87.0	10.8	2.1	(100.0)	88.2	9.2	2.5	(100.0)	87.6	10.0	2.4	(100.0)

*†As defined in section 3.3.3.

Figure 3.9. States of survivorship. Colon, rectum and anus cancer, UK, 2008.

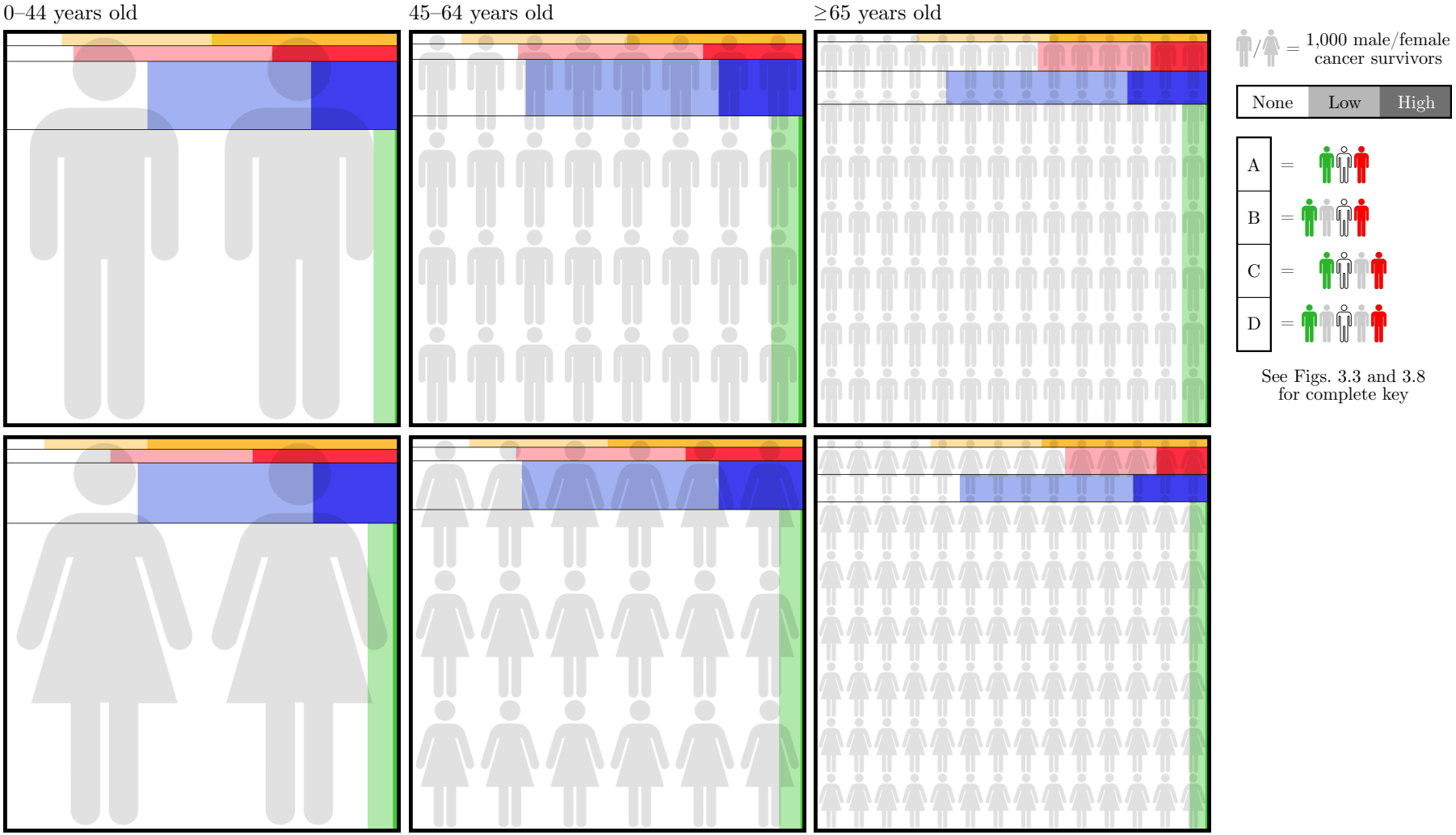


Figure 3.10. States of survivorship. Lung, bronchus and trachea cancer, UK, 2008.

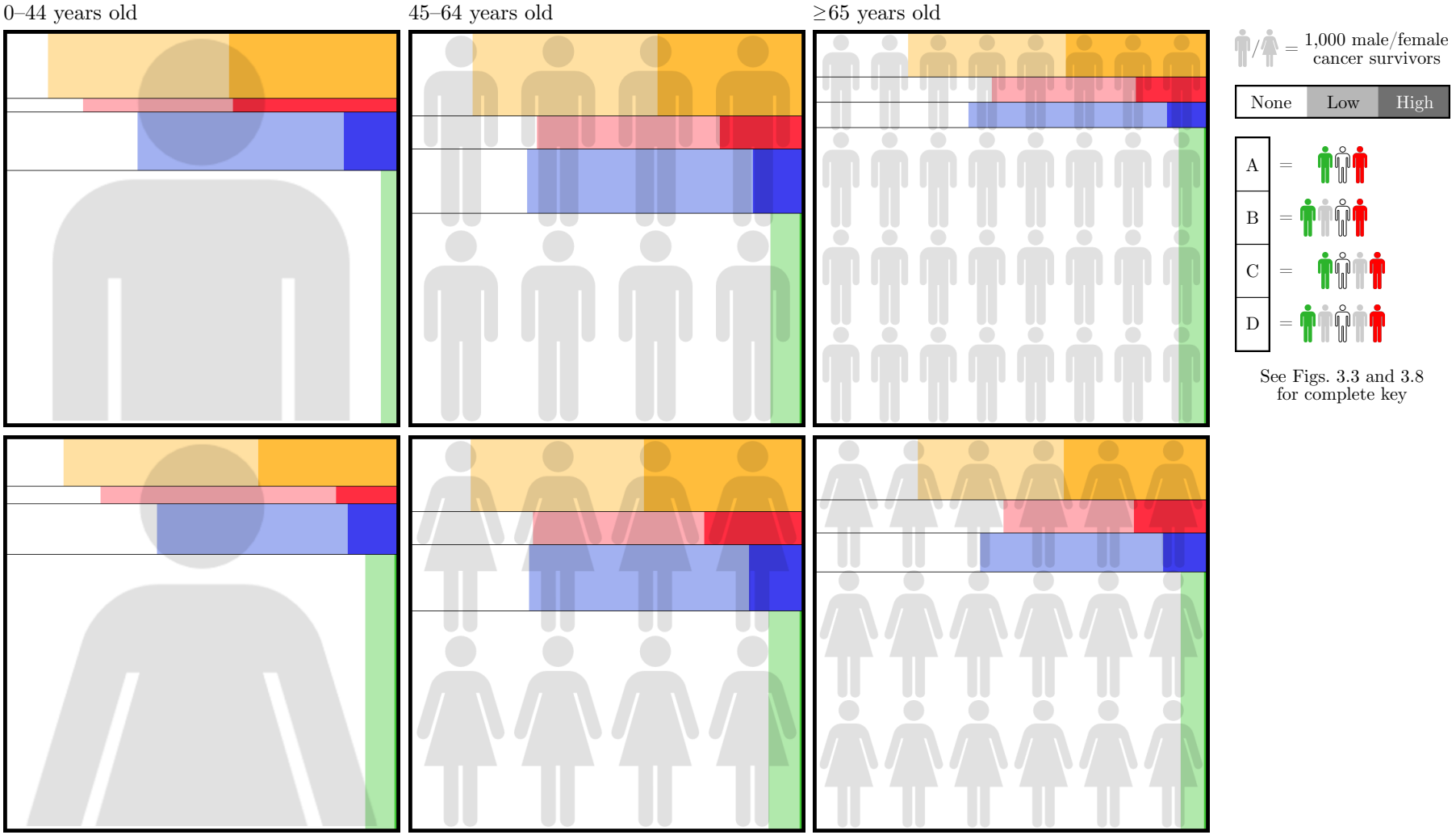
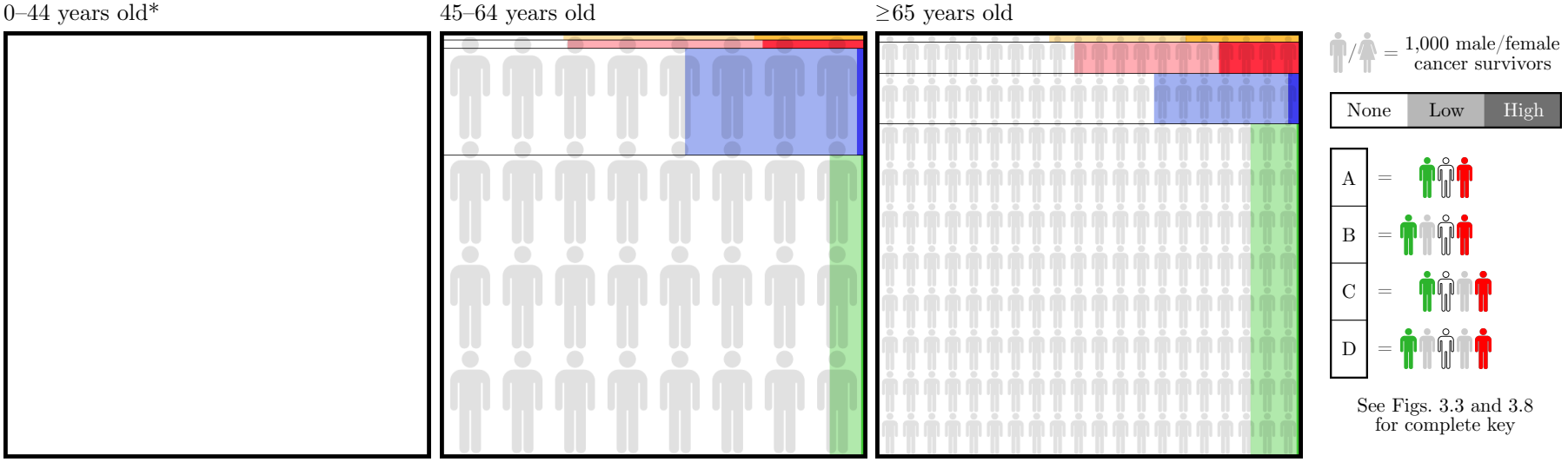


Figure 3.11. States of survivorship. Prostate cancer, UK, 2008.



*Display suppressed due to very small numbers in this age group.

Figure 3.12. States of survivorship. Female breast cancer, UK, 2008.

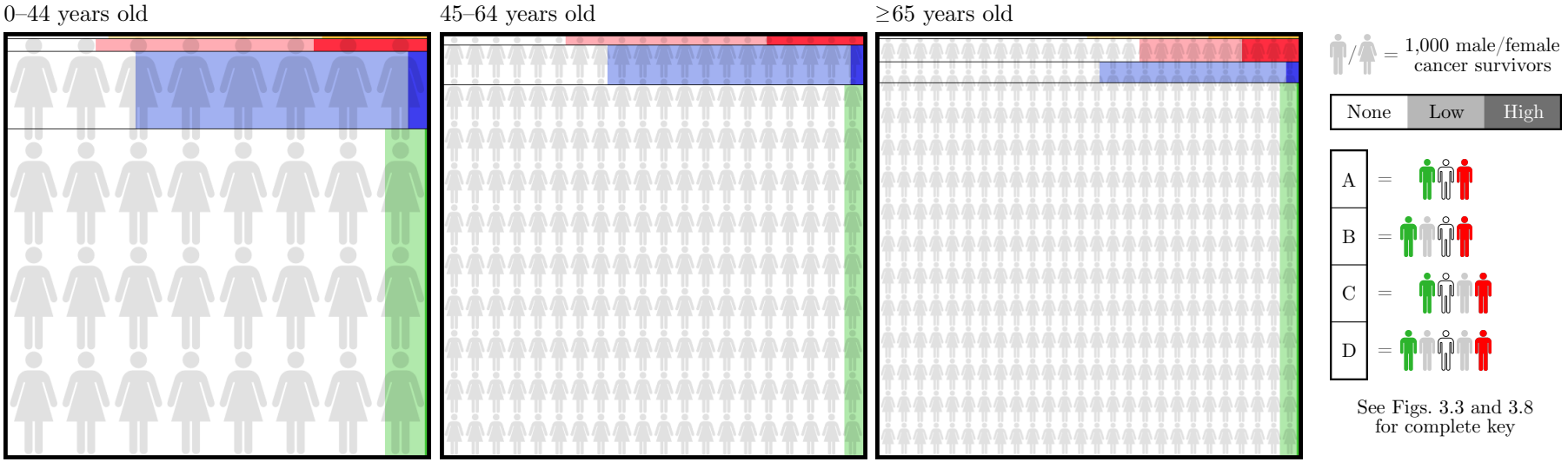
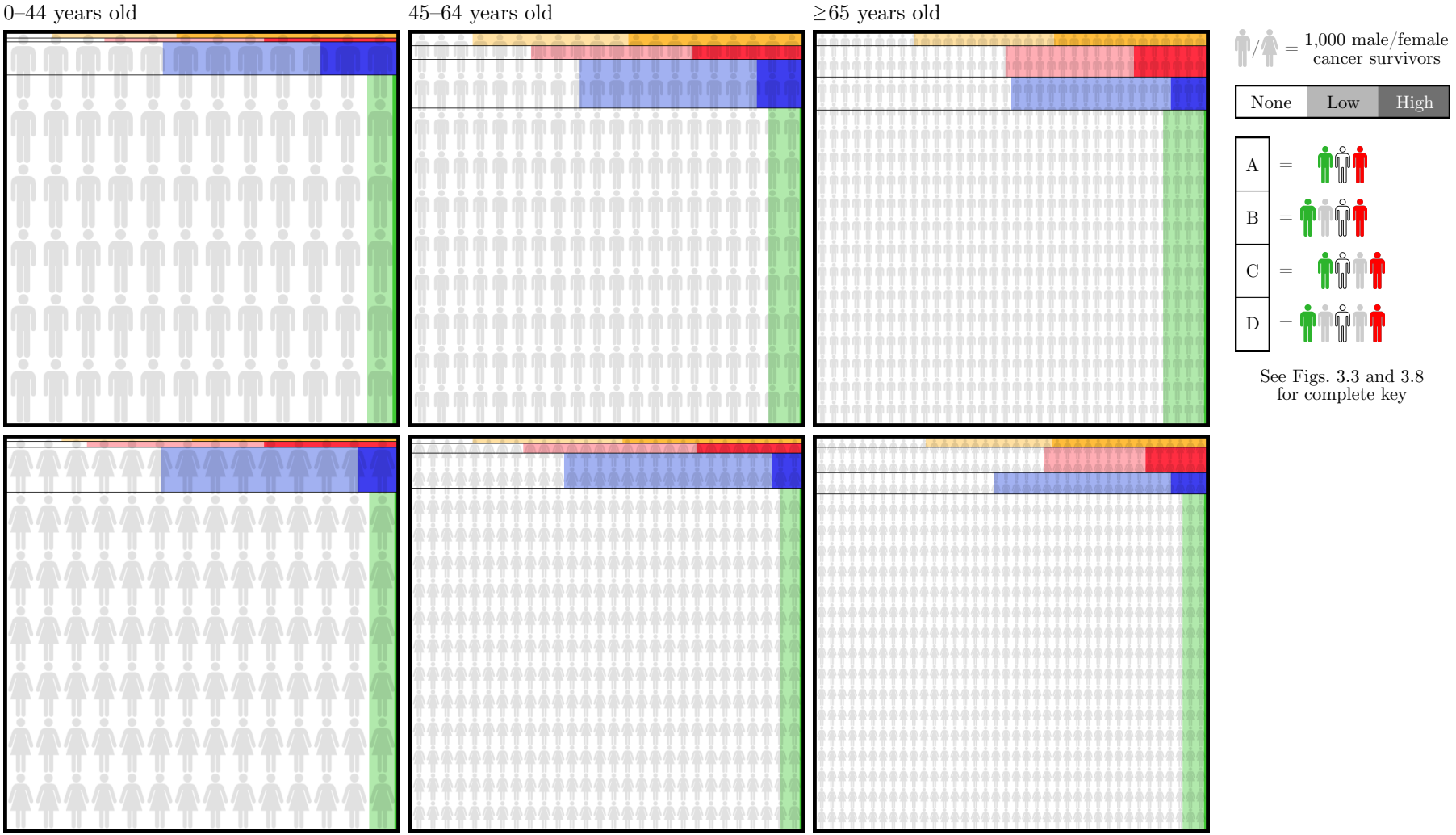


Figure 3.13. States of survivorship. All malignant neoplasms (excluding non-melanoma skin cancer), UK, 2008.



3.5 Discussion

In this chapter, a person-time analysis of linked cancer registry and hospital activity data for England was presented. Attained age, time since diagnosis and time until death were anticipated to be important factors in the level of health service utilisation among cancer survivors, and so the analysis was disaggregated according to these variables. The linked dataset allowed an analysis of all recorded episodes of NHS hospital in-patient or day case health care among registered cancer survivors in England during 2006. The person-time approach made it possible to quantify the intensity of hospitalisation by considering the proportion of time spent in hospital – this was considered to be a better measure of the cancer burden to both survivors and the health service than simply the number of hospital admissions.

There were two main strands to the analysis. Firstly, methods were developed to classify person-time of prevalence and person-time of hospital activity and to estimate the mean proportion of time spent admitted to hospital in the year 2006 by an individual sampled at random from the prevalent population, and how this varied with attained age and time since diagnosis. Secondly, different temporal phases of survivorship were identified and defined based on time since diagnosis and time until death. Time spent in these phases was described according to the intensity of cancer related acute health service utilisation that occurred, and from this the number of cancer survivors in the UK in different states of survivorship at the end of 2008 was estimated. It should be noted that an implication of this methodology is that survivorship states at a given point in time are influenced by hospital admissions both in the recent past and near future.

These two analysis stages are complementary. In the first stage, person-time was pooled for the whole population, allowing a precise assessment of the influence of attained age and time since diagnosis on the mean number of days (per 100 person-days) spent admitted to hospital by cancer survivors; but this approach provided a measure only of the *average* health service utilisation and not the underlying distribution. The second stage, however, explored this distribution by defining different categories of intensity of health service utilisation for individual survivors. It was shown that the underlying distribution of hospital activity is skewed heavily towards cancer survivors with no hospital activity, particularly in the period more than one year after diagnosis. This should be kept in mind when considering the mean proportion of time spent in hospital for the whole population.

The hospital activity data from HES featured details of all types of in-patient and day case care delivered to cancer survivors, not just that which related to cancer. Each record had up to 14 diagnostic codes (using the ICD-10 classification) and 12 operation procedure codes (using the OPCS-4 clinical classification), with the first of each of these codes intended to indicate the primary diagnosis or intent of the treatment. For simplicity, episodes were categorised as 'cancer related' or 'non-cancer related' according to the diagnostic codes only – an episode was considered to be cancer related if any of the diagnostic codes was between C00 and C97 (excluding C44). This was an intentionally broad categorisation, designed to take account of the wide range of health problems associated with cancer and the side effects of its treatment, as well as to negate any possible institutional variation in coding of the primary diagnosis. Defined thus, 'cancer related' hospital activity may be considered as that directly or indirectly caused by, or associated with, a cancer diagnosis. Nonetheless the specificity of the definition is unproven and it is not possible to say exactly which kinds of treatment following a cancer diagnosis are included – a specific analysis of the types of recorded clinical procedures was considered to be beyond the scope of this work due to the extremely large number of different operation codes contained in the data. However, analysis of the complementary 'non-cancer related' hospital activity (Figure 3.6) indicated that the cancer related definition achieved its goal of removing some of the background hospitalisation experienced by this population – the proportion of time spent by cancer survivors in hospital for non-cancer related care was much lower than for cancer related care, and, importantly, showed little variation with time since diagnosis.

Perhaps a more precise way of defining the health care burden directly attributable to cancer would be to compare hospitalisation of any kind among cancer survivors to that observed in a randomly selected age and sex-matched subset of the cancer-free population (or, more practically, the general population). This would allow some measure of the background hospitalisation experienced by the cancer-free or general population to be removed from the levels observed in cancer survivors. However, no such dataset was available for comparison – the Cancer Registries' National HES extract contains only episodes of care for patients with at least one episode 'for or with cancer' – and so the broad cancer/non-cancer related distinction described above was considered the most practical.

Some survivors will be diagnosed with additional primary cancers some time after their first diagnosis. Indeed, many studies have shown elevated cancer incidence rates in

those previously diagnosed with cancer compared with the general population (Evans et al., 2001, 2002) and the prevalence of multiple malignancies among cancer survivors has been shown to be around 7% in the Netherlands (Liu et al., 2011). However, treatment received for subsequent cancers is indistinguishable in this analysis from that received for the initial cancer, a fact to be remembered when considering the 'time since diagnosis' dimension. Furthermore, a small number of survivors in the cancer registry data will have had a diagnosis pre-1990 as well as in the period 1990–2006. However, since details of diagnoses made before 1990 were not available in the linked dataset, it was necessary to assume that all survivors in the dataset who were alive during 2006 had their first diagnosis in the period 1990–2006. The effect of this assumption is a possible small re-distribution of survivors between the types of cancer studied, but is not considered to be a significant limitation.

Each day of hospitalisation was treated equally in this analysis, and it is a limitation that the variation in the burden presented by different types of admission and different clinical procedures is not taken into account. For example, some survivors will be admitted to hospital for routine observation and monitoring, whereas others will be admitted (possibly in an emergency) for complicated operations that consume large amounts of hospital resources and require intense periods of rehabilitation.

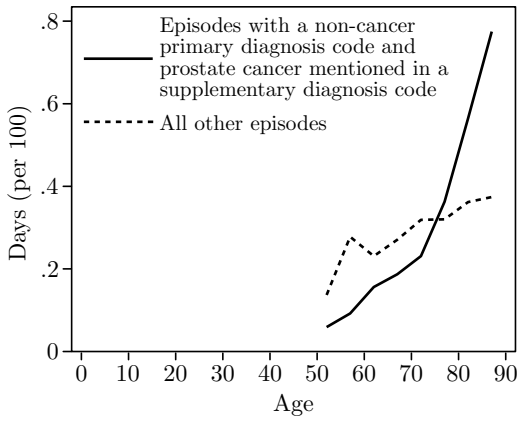
3.5.1 Hospital activity among cancer survivors

The majority of cancer related health service utilisation occurred during the first year after diagnosis, for survivors of all the cancer types studied. Most cancer patients receive some form of care or treatment as soon as possible after diagnosis, and thus it is perhaps unsurprising that this period contained a large amount of hospital activity. However, there was also a significant amount of cancer related health service utilisation in the period 1–5 years after diagnosis, particularly among survivors of lung and colorectal cancers. This is no longer the initial treatment phase, but is indicative of the ongoing care needs of cancer survivors.

The highest levels of cancer related acute health service utilisation were observed in survivors of the relatively poor prognosis cancers, but these differences largely disappeared more than five years after diagnosis. This indicates that the worse prognosis cancers required more intensive treatment regimes (including end of life care) in the short to medium term after diagnosis, but not, for those surviving, in the long term.

Prostate cancer survivors over the age of 60 and at least five years beyond diagnosis had higher levels of cancer related acute health service utilisation than comparable survivors of the other cancer types (see Figure 3.5 and Tables 3.2, 3.4, 3.6 and 3.8). This was despite prostate cancer survivors generally having relatively low levels of health service utilisation. A detailed investigation showed that these survivors experienced a large amount of hospital activity recorded in the HES data with a non-cancer primary diagnosis code, but with prostate cancer recorded as one of the supplementary diagnoses, and it was largely these episodes that resulted in the relatively high number of days spent admitted to hospital (per 100 person-days) among this group (Figure 3.14).

Figure 3.14. Mean cancer related admitted patient hospital activity among prostate cancer survivors five or more years after diagnosis, England, 2006. Number of days admitted to hospital per 100 person-days, by attained age.



In the analysis, these episodes were considered to be 'cancer related' because cancer was recorded as one of the supplementary diagnosis codes in the HES data. However, the recorded primary diagnosis codes, together with the operation procedure codes, indicated that these episodes were not generally related directly to prostate cancer. For example, the most common procedure codes were endoscopic examinations of the bladder and urethral catheterisations of the bladder, and the most common primary diagnosis codes were for disorders of the urinary system. There are several physiological changes that occur in men as they get older and lead to alterations in lower urinary tract function, making urinary disorders in elderly men common (Dubeau, 2006; Griebling, 2008). Many prostate cancer survivors are likely to be closely monitored in urology clinics for many years after initial diagnosis – often an extended period of either 'active surveillance' or 'watchful waiting' is pursued, especially if the cancer is less aggressive and treatment is not immediately necessary (Parker, 2004). This may

explain the greater hospital activity, many years after diagnosis, among prostate cancer survivors over the age of 60, compared with male lung or colorectal cancer survivors – it is possible that the increased observation and monitoring of these survivors leads to a high level of urological intervention which is only indirectly related to their prostate cancer diagnosis.

For all cancer types apart from female breast, acute health service utilisation in the first year after diagnosis was significantly higher for survivors aged over 70 compared with younger age groups. Initial cancer treatment can be physically very arduous, particularly for older patients who may be more frail and suffering a greater number of co-morbidities, resulting in more frequent and extended admissions to hospital. Particularly striking was the three-fold difference between cancer related acute health service utilisation in the first year after diagnosis among prostate cancer survivors aged at least 85 compared with those aged 65–69 (Figure 3.5). Since the early 1990s, the PSA test has increasingly been available (usually to men aged over 50) as a screening tool for prostate cancer, although in the UK no organised PSA screening programme is in place. This test is still controversial and considered to potentially result in over-diagnosis and over-treatment of prostate cancers which would otherwise never have become symptomatic (Barry, 2009; van Leeuwen et al., 2010). Recorded incidence rates and survival have, accordingly, greatly increased since the test's introduction (Evans and Møller, 2003; Cancer Research UK, 2010, 2011c), as have the number of prostate cancer survivors (Chapter 2). Men diagnosed with prostate cancer over the age of 70 are more likely to have symptomatic disease requiring intensive initial treatment compared with those aged 50–70 years who are more likely to have been diagnosed via the PSA test. Similarly, the observed reduction in acute health service utilisation in recently diagnosed female breast cancer survivors aged 50–70 (the age range in which women are routinely invited to attend breast screening units in England (NHS Cancer Screening Programmes, 2011)) reflects some of the benefits of early detection offered by screening programmes.

Non-cancer related acute health service utilisation was shown to generally increase with age, reflecting the greater number of morbidities likely to be found in the elderly compared with the young. Generally it varied little with time since diagnosis, but for lung and colorectal cancer survivors levels were slightly higher for shorter time since diagnosis. This may be explained by the fact that some survivors will have been diagnosed during an emergency admission to hospital. Also, some genuine cancer related hospital activity may not be coded as such in the HES data.

3.5.2 Survivorship states

As described in section 3.3.3, and illustrated in Figure 3.3, four mutually exclusive temporal phases of survivorship were defined according to time since diagnosis and time until death, as follows:

- Phase A: less than one year before death and less than one year after diagnosis.
- Phase B: less than one year before death and more than one year after diagnosis.
- Phase C: more than one year before death and less than one year after diagnosis.
- Phase D: more than one year before death and more than one year after diagnosis.

The first year following diagnosis is the time during which cancer patients receive initial treatment, the success of which may significantly affect their subsequent health and well-being. The final year before death is also significant since, for many people who die from cancer, there is a period of health deterioration in the months beforehand. For some survivors, post-diagnosis survival time is short enough that the first year following diagnosis and the final year of life overlap. Conversely, for those who live long enough after diagnosis, the period of survivorship which is more than one year after diagnosis and more than one year before death may be characterised by periods of remission, relapse, disease monitoring and/or eventual 'cure'. The temporal phases of survivorship defined above therefore provide a broad but informative way of describing the population of cancer survivors, and at any given point in time every survivor is in precisely one of these phases.

Lung cancer has a universally poor prognosis (age-standardised 5-year relative survival in England and Wales is under 10% (Cancer Research UK, 2009b)) and accordingly a large proportion of lung cancer survivors were less than one year beyond diagnosis and also less than one year from death (phase A). On the other hand, one quarter of prostate cancer survivors aged 45–64 were less than one year beyond diagnosis but not less than one year from death (phase C), reflecting the relatively good prognosis of this disease and the rapidly increasing incidence rates brought about by the use of the PSA test since the early 1990s (Evans and Møller, 2003).

In recent years, the number of cancer survivors in the UK has grown steadily each year (Chapter 2: Figure 2.3). The distribution of survivors between temporal phases, and the intensity of acute health service utilisation within them, provides an insight into what is meant by the term 'cancer survivor', especially given the current national survivorship initiatives in the UK and the movement towards understanding cancer as a chronic

illness. For example, the majority of UK cancer survivors (1.69 million of the 2.00 million as at 2008) are more than one year beyond diagnosis and more than one year from death, and the degree of acute health service utilisation in this phase is small – 1.57 million are in a period characterised by no cancer related acute health service utilisation and these survivors account for 78% of all UK cancer survivors.

Cancer survivors can now realistically expect to live longer beyond diagnosis, but this analysis suggests that the primary burden of cancer on the health service still comes from survivors in the first year following diagnosis and/or near the end of their life. The term 'cancer survivor' was originally proposed by the US National Coalition for Cancer Survivorship in 1986 at a time when "cancer was a disease that people needed to learn to fight" but has, according to some, become "so muddy that...a new definition is needed" (Twombly, 2004). Alternative terms include people 'living with or beyond cancer' or 'cancer patients'. However, this analysis has shown that the population of cancer survivors is heterogeneous, and finding a single term to usefully define everyone who has ever been diagnosed with cancer may not be possible.

A limitation of this work, brought about by the extent of the available hospital activity data, is that it only considered admitted hospital episodes of care (i.e. in-patients and day cases). Visits to general practitioner surgeries and other out-patient clinics are not captured in this analysis, but much of the observation and monitoring of survivors (especially those who are more than one year beyond diagnosis and more than one year from death) is carried out in such clinics (Khan et al., 2010). Neither does this analysis consider the personal psychosocial or general health burden of cancer on survivors – the trauma of being diagnosed with a life-threatening illness such as cancer is associated with post traumatic stress disorder, depression, and other mental disorders (Smith et al., 1999; Gregurek et al., 2010) and cancer survivors have been found to have poorer general health outcomes than individuals who have not been diagnosed with cancer (Yabroff et al., 2004). Cancer survivors are also likely to face day-to-day struggles (such as financial, emotional, relationship and employment difficulties) even if they have no need for treatment in hospital (Short et al., 2008; Foster et al., 2009). These issues present significant burdens to cancer survivors and should be kept in mind when considering the distribution of survivors between the states defined in this analysis.

3.6 Summary

The extent to which the population of cancer survivors in the UK is receiving care and treatment in hospital is central to understanding the burden of cancer on society. The

work described in this chapter provides a detailed characterisation of cancer survivorship in the UK which greatly enriches the basic prevalence estimates of Chapter 2. The findings contained here will be of interest to health service providers keen to quantify the volume of acute health care administered to cancer survivors, and the associated financial burden, as well as to survivors themselves.

The previous two chapters focused on current cancer prevalence in the UK. What follows in the next three chapters is the derivation and application of a model for projecting cancer prevalence into the future.

Chapter 4. A model for projecting cancer prevalence

In the following three chapters a discrete time model of the dynamics of cancer prevalence is developed, tested and used to project cancer prevalence in the UK up to 2040 under various scenarios of future incidence and survival.

The model is introduced in this chapter and described by a set of mathematical equations that allow continuous time data (from cancer registry sources) to be used as the inputs. A computer program implementation of the model is then briefly described and its potential applications discussed. Although the primary motivation is to provide projections of cancer prevalence, the model is generic enough to be applied to any disease which is considered to be prevalent in all individuals from the moment they are diagnosed until death.

Exercises were carried out to test the model and these are described, together with the extent of the required input data, in Chapter 5. Finally, results showing projections of cancer prevalence in the UK for the next 30 years are contained in Chapter 6.

4.1 Introduction

Projections of cancer prevalence can be used to estimate the future burden of cancer and inform the likely resources that will need to be allocated in order to meet the many and varied needs of the population of survivors. They provide valuable intelligence to health service resource planners as well as those responsible for providing care and support in the community to people affected by cancer and its treatment. Estimates of future cancer incidence and mortality are routinely produced by various bodies at local and national levels (NHS Scotland, 2004; Olsen et al., 2008; Parkin et al., 2008; Gatenby et al., 2011; Mistry et al., 2011; Thames Cancer Registry, 2011), but projections of cancer prevalence are less common. In part this may be due to the extra level of complexity involved in projecting prevalence compared with other vital statistics. For example, the definition of complete prevalence is such that data spanning many years are required in order to estimate it (see Chapter 2). Additionally, although it may be intuitively obvious that increasing incidence and/or increasing survival will result in increasing prevalence, the exact mechanisms by which this occurs, and the extent of the effect of each on the other, are not often investigated.

For these reasons, a flexible and general description of the dynamics of cancer prevalence is sought in order to produce projections of prevalence and to investigate the

influence of incidence rates, survival and demographic changes on future prevalence in the UK.

4.2 Background and literature

Cancer prevalence is a function of cancer incidence and survival, since it increases as new diagnoses are made and decreases as people previously diagnosed die. Consequently, the task of projecting cancer prevalence consists of two exercises: a) defining appropriate equations which relate cancer prevalence to incidence and survival; and b) producing the required projections of cancer incidence and survival on which future prevalence depends. The first exercise describes the mathematical model of prevalence that will be used and the second produces the input data for this model. A small number of different approaches to these exercises have been described in the literature. Various authors have differed in their choice of input data (for example, incidence and survival, incidence and mortality or survival and mortality) and how these are modelled but, essentially, the mathematical relationships between prevalence, incidence, survival and mortality are a matter of fact.

In places where cancer registration provides complete coverage of the population, and a long time series of data is available, estimates of current cancer prevalence can be made directly. It is then possible to obtain future prevalence estimates by considering the future survival of those currently prevalent as well as the future survival of those yet to be diagnosed (Phillips et al., 2002; Heinavaara and Hakulinen, 2006). However, if there are only sparse or incomplete cancer registry data available then this is not possible and both current and future cancer prevalence estimates must be entirely modelled (De Angelis et al., 1994; Tabata et al., 2008).

De Angelis et al. (1994) produced a computer program (called MIAMOD) which uses mortality and survival data to produce modelled estimates of incidence and prevalence in situations where these cannot be estimated directly using cancer registry data. Verdecchia et al. (2002) produced a similar program (PIAMOD) which *does* require cancer registry data and uses incidence and survival to calculate prevalence and mortality. Both of these programs are based on the equations relating disease incidence, survival, mortality and prevalence described by Verdecchia et al. (1989) and can be used to estimate current complete prevalence, as well as to provide medium to long-term projections.

In the literature, there are numerous examples of the application of these methods, particularly by investigators in Italy. Capocaccia et al. (1995) used mortality and survival data to estimate the prevalence of stomach cancer in Italy in the period 1970–1990, and to project this to 2000. National mortality data were available, but survival was estimated based on cancer registry data from four Italian provinces. Despite decreasing incidence rates, stomach cancer prevalence was shown to have increased in the period 1980–1990 and was anticipated to continue increasing up to the year 2000. The same authors also produced a similar paper that focussed on colorectal cancer (Capocaccia et al., 1997); again, national mortality and regional survival data were used to describe and project cancer prevalence trends in Italy from 1970 to 2000. The number of colorectal cancer survivors was anticipated to double between 1990 and 2000 due, mainly, to rising cancer incidence and population ageing. These results were used to highlight the increasing burden presented by colorectal cancer to the health service in Italy, in terms of both short-term intensive care and long-term care aimed at preventing disease recurrence.

Verdecchia et al. (2001) used national estimates of cancer mortality, survival estimates from the ITACARE study (Verdecchia et al., 1997) and the MIAMOD program to project the prevalence of all cancers combined in Italy from 1990 to 2000. It was shown that cancer prevalence was increasing at a faster rate for males than for females. Overall, cancer prevalence was projected to be 70% higher in 2000 compared with 1970, and 23% higher compared with 1990. These dramatic increases were found to be largely due to population ageing. This work was extended by Grande et al. (2006) who projected the prevalence of all cancers combined up to the year 2010 at a regional level in Italy. These authors anticipated that the proportion of the population who were cancer survivors would increase by around 3.5% per annum between 1999 and 2010.

In the USA, Mariotto et al. (2006) used the PIAMOD program to project colorectal cancer prevalence from 2000 to 2020. Estimates of incidence and survival were made using data from the SEER group of regional cancer registries and, for the purposes of prevalence projection, assumed to be constant from 2000 to 2020. Under this assumption, the number of colorectal cancer survivors was projected to increase by 50% from 1.0 million to 1.5 million over the period. In Japan, Tabata et al. (2008) described a similar approach to that used in the PIAMOD program and produced estimates of national cancer prevalence using incidence and survival data from selected regional cancer registries. These authors estimated that national 5-year cancer prevalence would increase by around 70% from 1.3 million in 1995 to 2.2 million in 2020.

When building a model for cancer prevalence, many authors have parameterised cancer incidence using age-period-cohort models (De Angelis et al., 1994; Phillips et al., 2002; Verdecchia et al., 2002; Heinavaara and Hakulinen, 2006; Tabata et al., 2008). These can be specified in numerous ways but, generally, incidence rates are assumed to be a function of attained age and birth cohort with or without an additional calendar time effect.

Survival from cancer and survival from other causes of death are often considered separately in models of cancer prevalence. Relative survival has been modelled using cure mixture models (Phillips et al., 2002; Verdecchia et al., 2002; Heinavaara and Hakulinen, 2006; Tabata et al., 2008). These assume that only a proportion of those diagnosed with cancer experience an excess risk of death attributable to cancer, whereas the rest (the cured proportion) do not – survival of the non-cured proportion is then assumed to follow the Weibull distribution (Mudholkar and Srivastava, 1993). Heinavaara and Hakulinen (2006) developed this approach further by allowing the cured proportion to experience different non-cancer risks of dying compared with the general population.

Before any of the previously described models for projecting cancer prevalence can be used, projections of the input data must be made. Age-period-cohort regression models of incidence are well-suited to medium-term projections since age and cohort effects can reasonably be assumed to persist into the near future. However, likely future trends in cancer survival have been treated slightly differently. Some authors have suggested considering two scenarios: one in which temporal trends in cancer survival are assumed to continue into the future, and another in which cancer survival is assumed to remain constant into the future (Capocaccia et al., 1995, 1997; Verdecchia et al., 2002; Tabata et al., 2008). The former may be considered the 'optimistic' scenario, since cancer survival is generally increasing in Western countries, and the latter as 'pessimistic' or 'conservative'. Heinavaara and Hakulinen (2006) produced projections of cancer prevalence using such scenarios for both survival *and* incidence.

In the following section, a model for projecting cancer prevalence is described. This was developed in the context of national cancer registry data with a long time series being available. Current prevalence, estimated directly from these data (see Chapter 2), forms the basis of the model.

4.3 Model description

In this section, a discrete time model to describe prevalence as a function of disease incidence and mortality is derived. The approach is similar to that used by Fiorentino et al. (2011) but differs in the choice of mortality, rather than survival, as an input – a choice made to reflect the nature of the available input data. In addition, the variables age at diagnosis and time of diagnosis, used by Fiorentino et al., are replaced by attained age at the index date and time since diagnosis, respectively.

The derivation of the model assumes that prevalence, according to attained age and time since diagnosis, is known for a given point in time, the goal being to express prevalence at some future point in time as a function of this known prevalence, future disease incidence and future mortality.

4.3.1 Discrete time definitions

The prevalence model is based in discrete time, i.e. variables such as age, time and time since diagnosis may only take whole number values (0,1,2,... etc.). It is sufficient to use a time resolution of one year in what follows, but other resolutions are also possible; for example, one month or one day. The practical implications of this, and methods for obtaining the required discrete time input data, are discussed after the derivation.

The following discrete time definitions of the model variables are made:

4.3.1.1 Inflow

Let the inflow, $N_{T,a}$, be the number of newly diagnosed people who, in the model, enter the prevalent population at time T and are aged a at time T. Inflow is the discrete time equivalent of incidence (count) in the continuous time setting.

4.3.1.2 Prevalence

Let $V_{T,y,a}$ be the number of people alive aged a at time T who entered the prevalent population at time $T-y$, for $y \geq 0$ and $a \geq y$. The variable y is therefore referred to as the time since entry, and is the discrete time equivalent of time since diagnosis in the continuous time setting. When considering limited duration prevalence there is an additional restriction of $y \leq Y$, for some $Y \geq 0$. Since $V_{T,y,a}$ is defined for $y \geq 0$, people entering the prevalent population at time T are included in the prevalence count at time T such that $V_{T,0,a} = N_{T,a}$.

4.3.1.3 Mortality

Let $M_{T,y,a}$ be the probability of dying between time T and time $T+1$, for people alive aged a at time T who entered the prevalent population at time $T-y$ for $y \geq 0$ and $a \geq y$. Note that both $V_{T,y,a}$ and $M_{T,y,a}$ are defined only for $a \geq y$ since nobody may enter the prevalent population before they are born, but the situation where age is zero upon entry ($a = y$) is permitted.

4.3.2 Derivation of the model equations

Suppose that prevalence at time T_1 is known. The principle of the model is that the prevalent population at time T_2 (where $T_2 > T_1$) is the union of two sub-populations: a) the proportion of the prevalent population at time T_1 that survives to at least time T_2, and b) the proportion of those entering the prevalent population between times T_1 and T_2 that survives to at least time T_2.

Sub-population (a) is estimated using the sets of prevalence and mortality variables $\{V_{T_1,y,a} : a \geq y \geq 0\}$ and $\{M_{T,y,a} : T_1 \leq T < T_2, a \geq y \geq 0\}$, respectively, and (b) is estimated using the sets of inflow and mortality variables $\{N_{T,a} : T_1 < T \leq T_2\}$ and $\{M_{T,y,a} : T_1 \leq T < T_2, a \geq y \geq 0\}$, respectively, as follows.

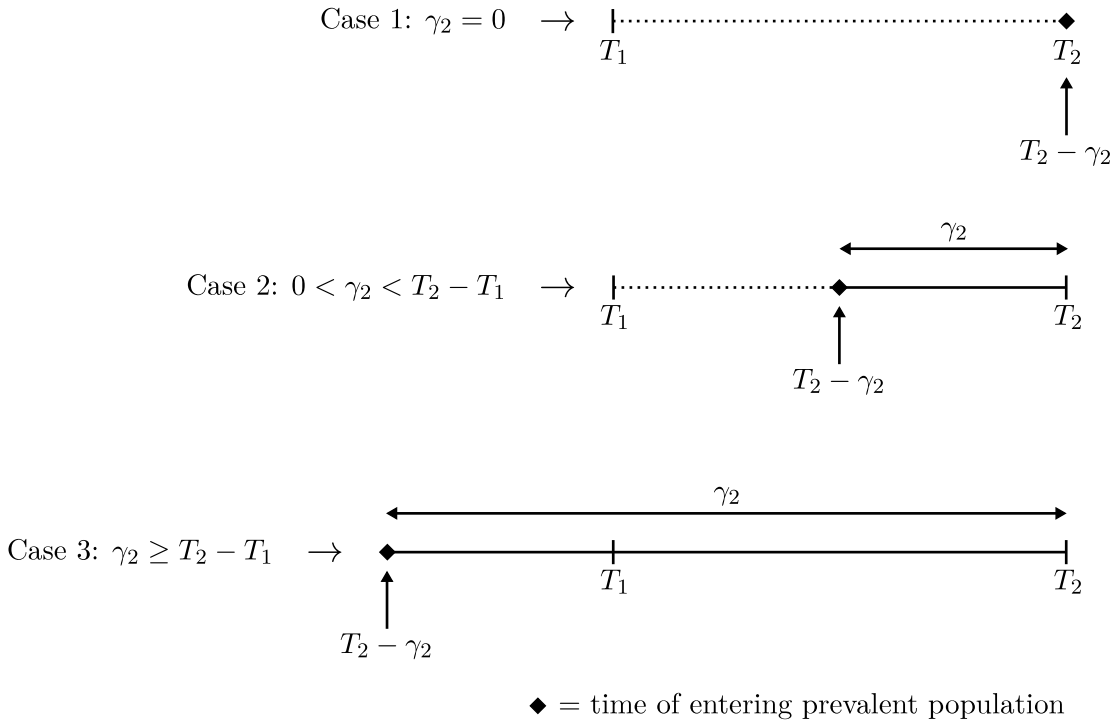
The probability that an individual who is prevalent at time T, aged a at time T having entered the prevalent population at time $T-y$, will survive from time T to at least time $T+1$ is given by $1 - M_{T,y,a}$. Therefore the probability of such an individual surviving from time T to at least time $T+I$, for some integer interval $I > 0$, is given by:

$$\prod_{i=0}^{I-1} \left(1 - M_{T+i,y+i,a+i}\right), \quad \forall a \geq y \geq 0 \qquad [4.1]$$

since, for every year that passes, age and time since entry to the prevalent population increase by one year, and the probability of surviving the entire interval is given by the product of the probabilities of surviving each consecutive year.

Using [4.1], and the principle stated above, an equation for prevalence $V_{T_2,\gamma_2,\alpha_2}$, the number of people alive aged α_2 at time T_2 who entered the prevalent population at time $T_2 - \gamma_2$, can be written down by considering the following three cases separately (Figure 4.1).

Figure 4.1. Timelines showing three different cases of time since entry to the prevalent population.



Case 1: $\gamma_2 = 0$. Then $V_{T_2,\gamma_2,\alpha_2}$ is simply the number of new patients who entered the prevalent population at time T_2, aged α_2. Therefore

$$V_{T_2,0,\alpha_2} = N_{T_2,\alpha_2}, \quad \forall \alpha_2 \geq 0\,. \qquad [4.2]$$

Case 2: $0 < \gamma_2 < T_2 - T_1$. Then $V_{T_2,\gamma_2,\alpha_2}$ is the number of people who a) entered the prevalent population at time $T_2 - \gamma_2$, b) were aged $\alpha_2 - \gamma_2$ at time $T_2 - \gamma_2$ and c) survived for a time interval of length at least γ_2 (i.e. from time $T_2 - \gamma_2$ to at least time T_2). Therefore, substituting $I = \gamma_2$, $T = T_2 - \gamma_2$, $y = 0$ and $a = \alpha_2 - \gamma_2$ into [4.1] gives:

$$V_{T_2,\gamma_2,\alpha_2} = N_{T_2-\gamma_2,\alpha_2-\gamma_2} \prod_{i=0}^{\gamma_2-1} \left(1 - M_{T_2-\gamma_2+i,i,\alpha_2-\gamma_2+i}\right), \quad \forall \alpha_2 \geq \gamma_2 \text{ if } 0 < \gamma_2 < T_2 - T_1\,. \qquad [4.3]$$

Case 3: $\gamma_2 \geq T_2 - T_1$. Here $V_{T_2,\gamma_2,\alpha_2}$ is a proportion of the known prevalence $V_{T_1,\gamma_1,\alpha_1}$, where $\gamma_1 = \gamma_2 - (T_2 - T_1)$ and $\alpha_1 = \alpha_2 - (T_2 - T_1)$. More precisely, $V_{T_2,\gamma_2,\alpha_2}$ is equal to $V_{T_1,\gamma_1,\alpha_1}$ multiplied by the appropriate probability of surviving for a time interval of length at least $T_2 - T_1$, i.e. from time T_1 to at least time T_2. Therefore, substituting $I = T_2 - T_1$, $T = T_1$, $y = \gamma_1$ and $a = \alpha_1$ into [4.1] gives:

$$V_{T_2,\gamma_2,\alpha_2} = V_{T_1,\gamma_1,\alpha_1} \prod_{i=0}^{T_2-T_1-1} (1 - M_{T_1+i,\gamma_1+i,\alpha_1+i}), \quad \forall \alpha_2 \geq \gamma_2 \text{ if } \gamma_2 \geq T_2 - T_1 . \tag{4.4}$$

Substituting for γ_1 and α_1, and altering the index i accordingly, gives:

$$V_{T_2,\gamma_2,\alpha_2} = V_{T_1,\gamma_2-T_2+T_1,\alpha_2-T_2+T_1} \prod_{i=\gamma_2-T_2+T_1}^{\gamma_2-1} (1-M_{T_2-\gamma_2+i,i,\alpha_2-\gamma_2+i}), \quad \forall \alpha_2 \geq \gamma_2 \text{ if } \gamma_2 \geq T_2 - T_1 . \tag{4.5}$$

So, combining [4.2], [4.3] and [4.5], the required expression for prevalence at time T_2, according to time since entry to the prevalent population γ_2 and age α_2 at time T_2, is given by:

$$V_{T_2,\gamma_2,\alpha_2} = \begin{cases} N_{T_2,\alpha_2} & \text{if } \gamma_2 = 0 \\ N_{T_2-\gamma_2,\alpha_2-\gamma_2} \prod_{i=0}^{\gamma_2-1} (1 - M_{T_2-\gamma_2+i,i,\alpha_2-\gamma_2+i}) & \text{if } 0 < \gamma_2 < T_2 - T_1 \\ V_{T_1,\gamma_2-T_2+T_1,\alpha_2-T_2+T_1} & \\ \quad \times \prod_{i=\gamma_2-T_2+T_1}^{\gamma_2-1} (1 - M_{T_2-\gamma_2+i,i,\alpha_2-\gamma_2+i}) & \text{if } \gamma_2 \geq T_2 - T_1. \end{cases} \tag{4.6}$$

4.3.3 Limited duration and complete prevalence

Limited duration $(Y+1)$-year prevalence at age α_2 is given by summing equation [4.6] over all permissible values of γ_2 up to the required maximum value, $Y \geq 0$, as follows:

$$\sum_{\gamma_2=0}^{\min(\alpha_2,Y)} V_{T_2,\gamma_2,\alpha_2} . \tag{4.7}$$

Furthermore, limited duration $(Y+1)$-year prevalence in a given age group $[A_L, A_U]$ can be calculated by summing $V_{T_2,\gamma_2,\alpha_2}$ over γ_2 and α_2 as follows:

$$\sum_{\alpha_2=A_L}^{A_U} \left(\sum_{\gamma_2=0}^{\min(\alpha_2,Y)} V_{T_2,\gamma_2,\alpha_2} \right) . \tag{4.8}$$

As Y increases, the above two summations approach complete prevalence at age α_2 and in age group $[A_L, A_U]$, respectively; i.e. that which includes everyone ever diagnosed.

It is also possible to segment limited duration prevalence at time T_2 into a sum of the contributions from the prevalent population at time T_1 and new entrants to the population after time T_1 (as in Fiorentino et al., 2011) as follows:

$$A_{T_1}(T_2) = \sum_{\gamma_2=0}^{\min(Y,T_2-T_1-1)} \left(\sum_{\alpha_2 \geq \gamma_2} V_{T_2,\gamma_2,\alpha_2} \right) \quad [4.9]$$

$$B_{T_1}(T_2) = \begin{cases} \sum_{\gamma_2=T_2-T_1}^{Y} \left(\sum_{\alpha_2 \geq \gamma_2} V_{T_2,\gamma_2,\alpha_2} \right) & \text{if } Y \geq T_2 - T_1 \\ 0 & \text{otherwise} \end{cases} \quad [4.10]$$

where, following the notation used by Fiorentino et al., $A_{T_1}(T_2)$ and $B_{T_1}(T_2)$ are the prevalence at time T_2 of those who enter the prevalent population after and before (or at) time T_1, respectively.

4.4 Input data

In order to use the model to estimate future disease prevalence, it is necessary to have a certain amount of input data. Cancer registry data are routinely used to estimate current and historical cancer incidence and prevalence, but it is also possible to use such data to estimate the required mortality probabilities for cancer survivors.

The model for projecting prevalence, as described above, was developed in the context of having a long time series of cancer registry data available, but the required input data may also be gathered from other sources. Either way, since the available data will only cover *historical* incidence and mortality, it is necessary to project these forward in time before using them as input to the model. Methods of projecting incidence and mortality are covered in Chapter 5.

The model is specified in discrete time, yet real world data are continuous; for example, inflow is defined as the discrete time version of incidence – the number of people entering the prevalent population at a discrete time point – but incidence in continuous time is usually reported over the course of a whole year and not all of those who are diagnosed in that year will survive long enough to be counted as prevalent at the end of the year. So it is necessary to consider in detail the translation of continuous time data into the required discrete time data.

4.4.1 Specification of discrete data inputs

The most convenient discrete time interval to consider is one year, and so in the following section the term 'at time T' is intended to mean 'at the first instant of year T'. In the discrete time model, at each time point T prevalence increases by $\sum_a N_{T,a}$ and simultaneously decreases by the number of people who were prevalent at time $T-1$ but who died between time $T-1$ and time T. It is assumed that all these deaths, as well as the inflow of newly diagnosed people, occur at precisely the same instant, time T. The prevalence count $V_{T,y,a}$ incorporates the resultant net flow of survivors into or out of the prevalent population at time T.

4.4.1.1 Incidence and inflow

In the real world, new diagnoses, which increase the size of the prevalent population, may occur at any time between two discrete time points; i.e. the inflow to the prevalent population at time T is made up of people who were diagnosed at any point in the year leading up to time T. However, not everyone diagnosed in this period will contribute to the prevalence count at time T since not everyone will survive to time T. For this reason, when specifying the model input data, $N_{T,a}$ should not be confused with the continuous time incidence in the year leading up to time T. Rather, $N_{T,a}$ should be estimated as the expected number of people diagnosed in the year leading up to time T who survive to time T and are aged a at time T, as follows:

Let $\widetilde{N}_{T,a}$ be the continuous time one year incidence count at age a in year T – i.e. the total number of people diagnosed with cancer at age a between discrete time points T and $T+1$. Then $N_{T,a}$ is comprised of a proportion of $\widetilde{N}_{T-1,a}$ and a proportion of $\widetilde{N}_{T-1,a-1}$, since everyone who enters the prevalent population at time T and is aged a at time T must have been diagnosed in the previous year at age a or $a-1$, depending on when in the year their birthday and date of diagnosis fall.

Let $\widetilde{N}^i_{T,a}$ be the number of people diagnosed with cancer at age a on the i^{th} day of year T. It is assumed that diagnoses are distributed equally throughout the year, so that

$$\widetilde{N}^i_{T,a} = \frac{\widetilde{N}_{T,a}}{\tau} \qquad [4.11]$$

where τ is the number of days in year T (i.e. 365 or 366; we may assume $\tau = 365$).

Let $\rho_{T,a,j}$ be the probability of dying on the j^{th} day after diagnosis for those diagnosed at age a in year T.

Consider those who are diagnosed on the i^{th} day of year $T-1$. For this group,

$$\Pr\left(\text{alive at } T \mid \text{diagnosed at age } a \text{ in year } T-1\right) = \prod_{j=0}^{\tau-i}\left(1-\rho_{T-1,a,j}\right) \qquad [4.12]$$

and, assuming that birthdays are distributed equally throughout the year,

$$\Pr\left(\text{aged } a \text{ at } T \mid \text{diagnosed at age } a \text{ in year } T-1\right) = \frac{i}{\tau} \qquad [4.13]$$

since each person will not age between diagnosis and time T if and only if their birthday falls earlier in the year than their date of diagnosis.

Therefore, the contribution to $N_{T,a}$ from $\widetilde{N}^i_{T-1,a}$ is given by:

$$\begin{aligned}&\widetilde{N}^i_{T-1,a} \times \Pr\left(\text{aged } a \text{ at } T \mid \text{diagnosed at age } a \text{ in year } T-1\right)\\ &\times \Pr\left(\text{alive at } T \mid \text{diagnosed at age } a \text{ in year } T-1\right)\\ =\ &\widetilde{N}^i_{T-1,a}\,\frac{i}{\tau}\prod_{j=0}^{\tau-i}\left(1-\rho_{T-1,a,j}\right).\end{aligned} \qquad [4.14]$$

Similarly, the contribution to $N_{T,a}$ from $\widetilde{N}^i_{T-1,a-1}$ is given by:

$$\begin{aligned}&\widetilde{N}^i_{T-1,a-1} \times \Pr\left(\text{aged } a \text{ at } T \mid \text{diagnosed at age } a-1 \text{ in year } T-1\right)\\ &\times \Pr\left(\text{alive at } T \mid \text{diagnosed at age } a-1 \text{ in year } T-1\right)\\ =\ &\widetilde{N}^i_{T-1,a-1}\left(1-\frac{i}{\tau}\right)\prod_{j=0}^{\tau-i}\left(1-\rho_{T-1,a-1,j}\right).\end{aligned} \qquad [4.15]$$

Summing [4.14] and [4.15] over i gives $N_{T,a}$, so that:

$$N_{T,a} = \sum_{i=1}^{\tau}\left(\widetilde{N}^i_{T-1,a}\,\frac{i}{\tau}\prod_{j=0}^{\tau-i}\left(1-\rho_{T-1,a,j}\right) + \widetilde{N}^i_{T-1,a-1}\left(1-\frac{i}{\tau}\right)\prod_{j=0}^{\tau-i}\left(1-\rho_{T-1,a-1,j}\right)\right). \qquad [4.16]$$

Then, substituting [4.11] into [4.16], gives:

$$N_{T,a} = \frac{1}{\tau^2}\sum_{i=1}^{\tau}\left(\widetilde{N}_{T-1,a}\,i\prod_{j=0}^{\tau-i}\left(1-\rho_{T-1,a,j}\right) + \widetilde{N}_{T-1,a-1}\left(\tau-i\right)\prod_{j=0}^{\tau-i}\left(1-\rho_{T-1,a-1,j}\right)\right). \qquad [4.17]$$

Note that $V_{T,y,a}$ is defined for all $y \geq 0$ and, from [4.6], $V_{T,0,a} = N_{T,a}$. So, a continuous time interpretation of $N_{T,a}$ is 1-year prevalence at age a and time T.

Counts of the historical numbers of incident cases of cancer, from which future counts can be extrapolated, are easily obtained from cancer registry data and therefore widely available. Similarly, the daily probabilities of death, $\rho_{T,a,j}$, may be estimated from cancer registry data using person-time methods (as described in the next section). Regression procedures for extrapolating future incidence and mortality from historical data are presented in Chapter 5.

4.4.1.2 Mortality probabilities

There are two types of mortality that must be estimated to provide input to the prevalence model. Use of equation [4.6] requires estimates of $M_{T,y,a}$ – the probability that a member of the prevalent population who is aged a at time T, and who entered at time $T - y$, will die before time $T + 1$. Use of equation [4.17] requires estimates of $\rho_{T,a,j}$ – the probability of death on the j^{th} day after diagnosis for a person diagnosed in the period $[T, T + 1]$, aged a at diagnosis.

Let $p_{T,y,a}$ be the daily probability of death in the period $[T, T + 1]$ amongst those aged a who, in continuous time, were diagnosed in the year leading up to time $T - y$ and therefore, in the discrete model, entered the prevalent population at time $T - y$. The primary distinction between $\rho_{T,a,j}$ and $p_{T,y,a}$ is that the former is the daily probability of death in the calendar year of diagnosis and the latter is the daily probability of death in subsequent calendar years. Additionally, it is assumed that the latter is constant throughout the year; i.e. in each calendar year subsequent to the calendar year of diagnosis, the risk of dying is the same on every day. This means that $M_{T,y,a}$ can be estimated, as described by Mayfield (1961, 1975) and Trent and Rongstad (1974), by:

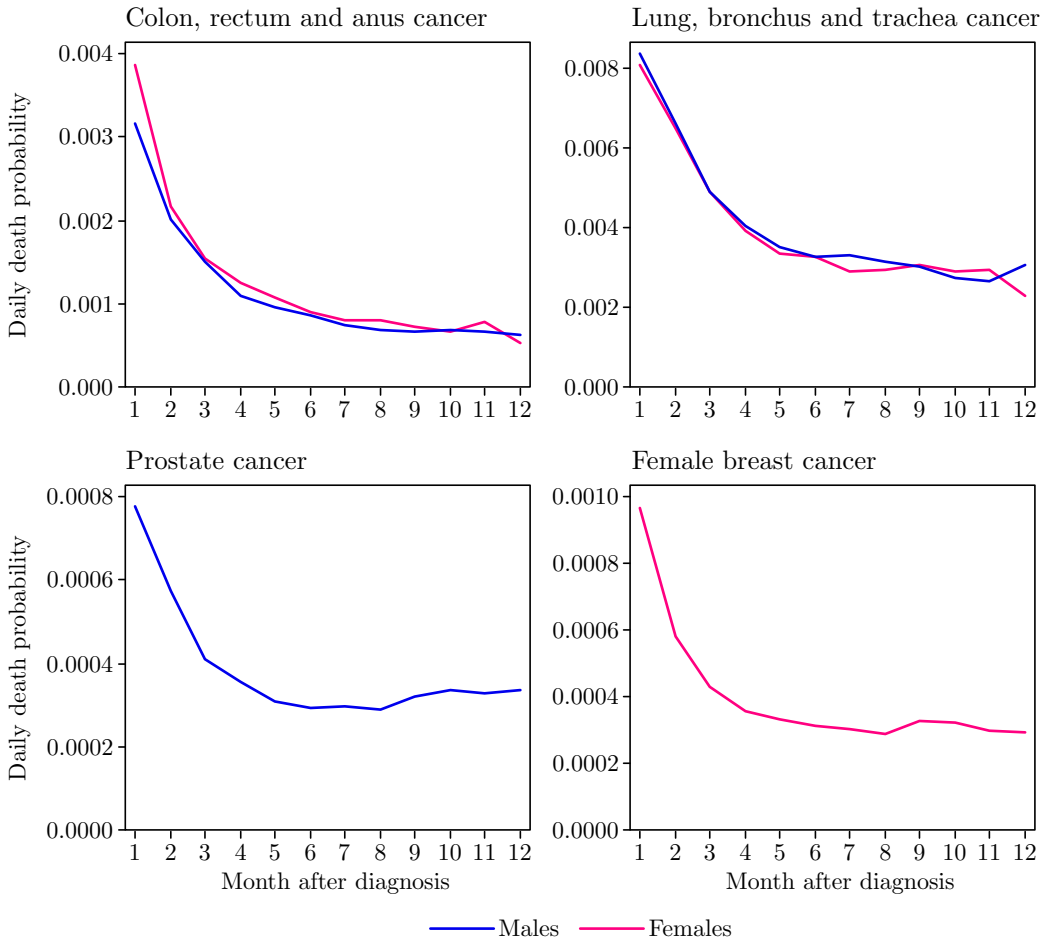
$$\widetilde{M}_{T,y,a} = 1 - (1 - p_{T,y,a})^{\tau} . \qquad [4.18]$$

A person-time analysis of cancer registry data was used to estimate the mortality probabilities $M_{T,y,a}$. In each year T and for each value of time since entry y and attained age a, the total number of deaths among survivors was calculated and divided by the appropriate total number of person-days of prevalence in the period to give an estimate of $p_{T,y,a}$, from which an estimate of $M_{T,y,a}$ was made using equation [4.18].

The assumption that, among cancer survivors, the daily probability of dying is constant throughout any given calendar year (but may vary between years) is used to estimate $M_{T,y,a}$, where the time since entry variable y is specified with a resolution of one year.

For cancer patients in England, it was found that the daily probability of death is much higher in the first few months following diagnosis than at other times, particularly among survivors aged 65 and over (see figure 4.2). Therefore, this assumption is not appropriate when estimating $\rho_{T,a,j}$ since it describes the daily probability of dying up to 12 months after diagnosis only and requires a time since diagnosis resolution of one day (via the index j).

Figure 4.2. Daily probability of death among cancer survivors in England aged 65 and over, 1999–2008, by time since diagnosis.



It is possible to estimate $\rho_{T,a,j}$ directly using cancer registry data and the same person-time method as for $p_{T,y,a}$ (described above). However, at the daily time since diagnosis resolution this method does not provide robust estimates since very small numbers are always involved. So instead, the daily mortality probabilities $\rho_{T,a,j}$ were first calculated assuming that they were constant on each post-diagnosis day in the calendar year of diagnosis. They were then scaled using weights to represent the archetypal dependence of daily mortality on time since diagnosis in the months following diagnosis, as follows:

Let $\rho_{T,a}$ be the daily post-diagnosis probability of dying for people diagnosed at age a in the time period $[T, T+1]$, under the assumption that daily post-diagnosis mortality is constant from the point of diagnosis up to time $T+1$; i.e. $\rho_{T,a}$ is equivalent to $\rho_{T,a,j}$ with the time since diagnosis dependence removed. Then $\rho_{T,a}$ can be estimated using cancer registry data and the same person-time methods that were used to estimate $p_{T,y,a}$ (described above).

Then let the weights $w_{j,a}$ be defined such that:

$$\begin{aligned}\prod_{j=0}^{\tau-i}(1-\rho_{T,a,j}) &= \prod_{j=0}^{\tau-i} w_{j,a}\,(1-\rho_{T,a})\,, \quad \forall T \text{ and } i \in [1,\tau] \\ &= (1-\rho_{T,a})^{\tau-i+1} \prod_{j=0}^{\tau-i} w_{j,a} \\ &= (1-\rho_{T,a})^{\tau-i+1}\, W_{i,a}, \quad \text{where } W_{i,a} := \prod_{j=0}^{\tau-i} w_{j,a}\,. \end{aligned} \qquad [4.19]$$

The role of these weights is to scale equation [4.17] to take account of the fine resolution dependence on time since diagnosis of the daily mortality probabilities in the first months following diagnosis that is not robustly estimable from the data. For simplicity, they were assumed to be independent of time T (as can be seen from the above definition), and were estimated using the most recent available 10 years of cancer registry data. It was also assumed that they were constant in broad age groups and in each month post-diagnosis (see section 5.3.2 for full details).

Then, equation [4.17] becomes:

$$\begin{aligned}N_{T,a} = \frac{1}{\tau^2} \sum_{i=1}^{\tau} \Big[&\widetilde{N}_{T-1,a}\; i\; (1-\rho_{T-1,a})^{\tau-i+1}\, W_{i,a} \\ &+ \widetilde{N}_{T-1,a-1}\,(\tau-i)\,(1-\rho_{T-1,a-1})^{\tau-i+1}\, W_{i,a-1}\Big]. \end{aligned} \qquad [4.20]$$

4.5 Implementation

A computer program was written using Microsoft Visual Basic for Applications (VBA) in the Excel computer package in order to execute the large number of calculations required by the model equations [4.6] and [4.20]. Each input dataset – current prevalence, future incidence, future mortality and the mortality weights described above – is copied by the user into a separate worksheet. When executed, the program reads in all the input data, performs the calculations, and places the results in the

output worksheets. The required input data were gathered from an analysis of cancer registry data, and regression models were developed where necessary (see Chapter 5 for full details).

4.6 Applications

The model is designed to be flexible and as general as possible. Estimates of future prevalence can be produced for any chosen input data – examples of this are given in Chapter 6. The only requirements are that full datasets are provided for current prevalence, future incidence, future mortality and future demographics. The data themselves may pertain to any geographical region, cancer type, sex, or anticipated future scenario.

One of the strengths of the model is in its ability to provide different prevalence estimates based on different scenarios relating to future incidence, mortality and demographics. The program may be run using future cancer statistics estimated from existing trends (perhaps incorporating various different assumptions) or alternatively using a set of pre-defined target statistics. This may be particularly useful to test the effect on prevalence of theoretical changes in survival. For example, cancer prevalence might be estimated under the assumption that cancer survival in the UK is to increase to a level comparable with that of another European population in the coming years.

However, the model equations require survival input data to be in the form of mortality probabilities for survivors, according to time since diagnosis and attained age, rather than survival explicitly. It would be useful to be able to specify scenarios in terms of changes in survival using standard survival analysis techniques, rather than changes in these mortality probabilities. Therefore, a method for translating changes in survival to changes in mortality probabilities for survivors is required.

4.6.1 Survival versus mortality

Proportional hazard models are often used to compare and analyse survival in two (or more) different populations. Such models assume that the two survival functions $S_1(t)$ and $S_2(t)$ are related by a constant hazard ratio HR. As shown in Collett (2003: p.47), this implies that

$$S_2(t) = S_1(t)^{HR}. \qquad [4.21]$$

Let M_1 and M_2 be the probabilities of dying between time t and time $t + \delta t$ under survival functions $S_1(t)$ and $S_2(t)$, respectively. Then

$$\begin{aligned} M_2 &= 1 - \frac{S_2\left(t + \delta t\right)}{S_2\left(t\right)} \\ &= 1 - \left(\frac{S_1\left(t + \delta t\right)}{S_1\left(t\right)}\right)^{HR} \\ &= 1 - \left(1 - M_1\right)^{HR}. \end{aligned} \qquad [4.22]$$

Equation [4.22] provides a method for translating changes in survival (under a proportional hazards model) to the equivalent changes in the required mortality inputs.

4.6.1.1 Example

Suppose that $S_1(t)$ represents the current cancer survival function for a particular cancer in the UK, and $S_2(t)$ represents a comparator or target survival function. The hazard ratio, HR, relating these survival functions can be estimated using standard proportional hazards regression techniques. The set of current mortality probabilities among survivors in the UK, $\{M_1\}$, is known and so the set of mortality probabilities $\{M_2\}$ can be estimated using [4.22]. These can then be used as input to the model for projecting cancer prevalence.

4.7 Summary

In this chapter the mathematical relationships between incidence, mortality and prevalence were set out and used to define a model for projecting cancer prevalence. Careful consideration was given to the task of specifying the model inputs based on cancer registry data, and a strategy for doing this was described. A brief discussion of the possible applications of the model was also included.

The focus of the following chapter is the model input data – the extent of data that are typically available in a cancer registry context is described and compared in detail to that which is required as input to the model. Methods for estimating future incidence and mortality, as required by the model, are developed and the performance of the model is then assessed. Finally, incidence and mortality projections for England up to the year 2039 are displayed. In Chapter 6, results showing projections of cancer prevalence in the UK up to the year 2040, under various different scenarios of future incidence and survival, are displayed and discussed.

Chapter 5. Projection model input data and evaluation

In this chapter, the input data required to execute the model for projecting cancer prevalence (described in Chapter 4) are explored. These data are described in terms of 'parameter spaces'. In this context, a parameter space is the set of parameter values for which a given input variable is required – for example, the ages and years for which it is necessary to provide cancer incidence counts. Parts of these parameter spaces (hereupon referred to simply as 'spaces') are readily available from an analysis of cancer registry data. The remaining parts must be estimated, and for this task a series of regression models are developed. Exercises designed to test the prevalence model and evaluate its forecasting ability are then described. Finally, projections of incidence and mortality up to the year 2039 are displayed.

5.1 The parameter spaces

The model for projecting cancer prevalence described in Chapter 4 was developed in the context of a large volume of cancer registry data being available, and depends upon a large amount of input data. These inputs are current prevalence counts, future incidence counts and future mortality probabilities for cancer survivors (see equations [4.6] and [4.20]).

Current prevalence may be counted directly using cancer registry data, with the only restriction being the length of the available data series. For example, in England cancer registration has been active since the early 1960s and a national quality assured dataset is available covering diagnoses made from 1971 onwards which, currently, allows estimates of almost 40-year limited duration prevalence to be made. Methods exist to adjust limited duration prevalence to account for survivors who were diagnosed before the start of the registry dataset (Capocaccia and De Angelis, 1997), but with such a long time series of data only small adjustments are needed (as in Chapter 2).

Future incidence counts and mortality probabilities must be estimated and this can be achieved by considering trends observed in the available data. Before designing appropriate methods for estimating future incidence and mortality, it is useful to consider in detail the space of available data and the space of data that is required to run the prevalence model.

5.1.1 Extent of available data

The notation of variables used in this chapter follows the conventions set out in Chapter 4. There are two variables that index incidence – namely, calendar time T and age a. The extent of the incidence data can therefore be visualised in a two-dimensional space (i.e. a plane) with axes $\{T, a\}$. Similarly, there are three variables that index cancer prevalence and survivor mortality probabilities – calendar time T, age a and time since diagnosis y – and therefore the extent of these data can be visualised in a three-dimensional space with axes $\{T, a, y\}$.

Suppose that the available cancer registry data cover all diagnoses made from year T_0 to year $T_1 - 1$ (inclusive) and that the maximum possible attainable age of members of the population is $a_{\max}$. Then the space of available incidence data (Figure 5.1) is given by:

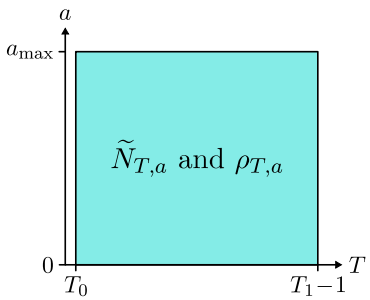
$$\left\{\widetilde{N}_{T,a} : T \in [T_0, T_1 - 1], a \in [0, a_{\max}]\right\} \qquad [5.1]$$

where T and a are discrete integer variables and $\widetilde{N}_{T,a}$ is the count of the number of cancers diagnosed in year T at age a.

The extent of the available daily mortality probabilities among survivors in the calendar year of diagnosis, $\rho_{T,a}$, is the same as for the incidence counts:

$$\{\rho_{T,a} : T \in [T_0, T_1 - 1], a \in [0, a_{\max}]\}\,. \qquad [5.2]$$

Figure 5.1. Extent of available incidence and mortality in the calendar year of diagnosis data.



Now let's consider the extent of the available prevalence, $V_{T,y,a}$. The cancer registry data contain a total of Y years of diagnoses, where Y is defined as $T_1 - T_0$. From these data, estimates of Y-year limited duration prevalence at the start of year T_1, given by $\sum_{y=0}^{Y-1} V_{T_1,y,a}$ in the discrete time formulation, are possible. In general, at an intermediate discrete time point $T \in [T_0 + 1, T_1]$, the restriction on y (the time since

entry to the prevalent population) in the available prevalence data is $y \in [0, T - T_0 - 1]$. Additionally, since no person may enter the prevalent population within the model before they are born, the age variable is restricted such that $a \geq y$. These conditions lead to the following space of available prevalence estimates:

$$\{V_{T,y,a} : T \in [T_0 + 1, T_1], y \in [0, T - T_0 - 1], a \in [y, a_{\max}]\} . \qquad [5.3]$$

Figure 5.2 shows cross sections of this space in each of the planes $\{T, a\}$, $\{y, a\}$ and $\{T, y\}$. In the three dimensional space $\{T, a, y\}$, the extent of the data forms a shape similar to a wedge of cheese but with a tapered underside (Figure 5.3).

Figure 5.2. Cross-sectional planar extents of available prevalence data.

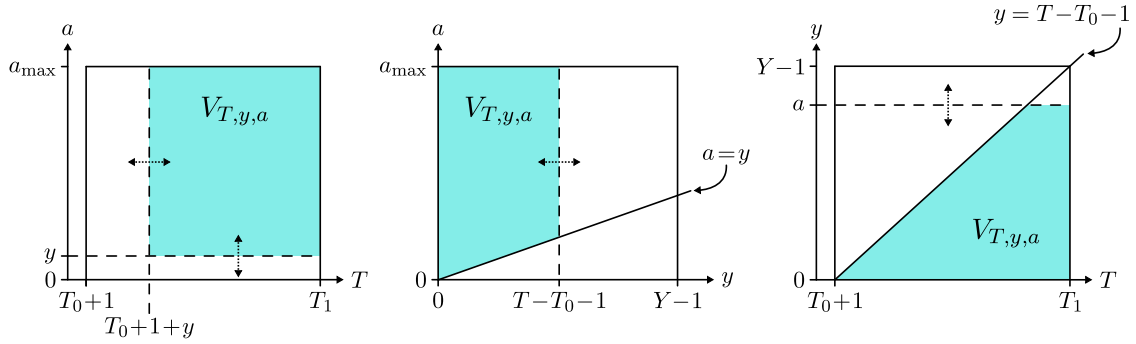
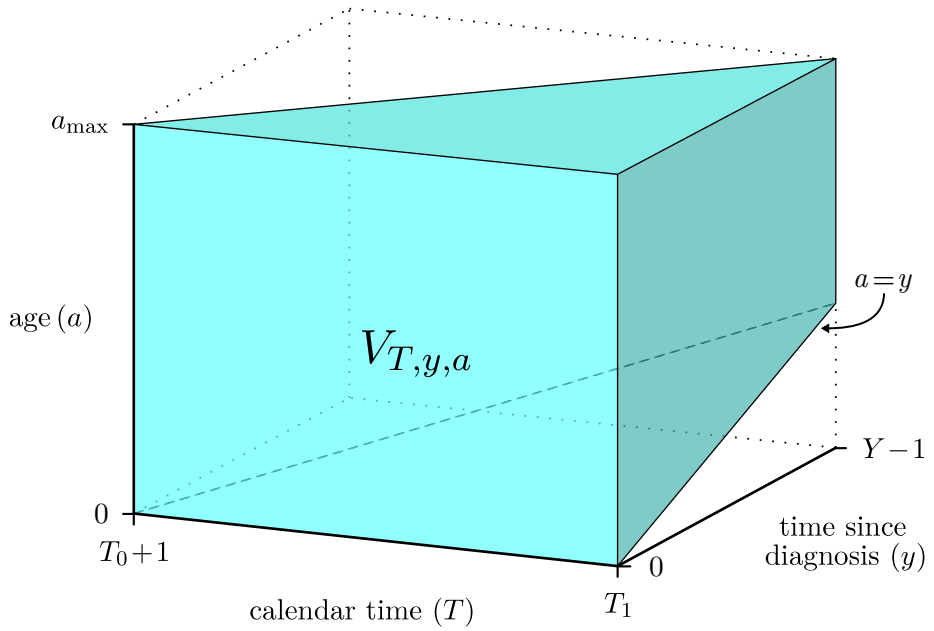


Figure 5.3. Three-dimensional extent of available prevalence data.

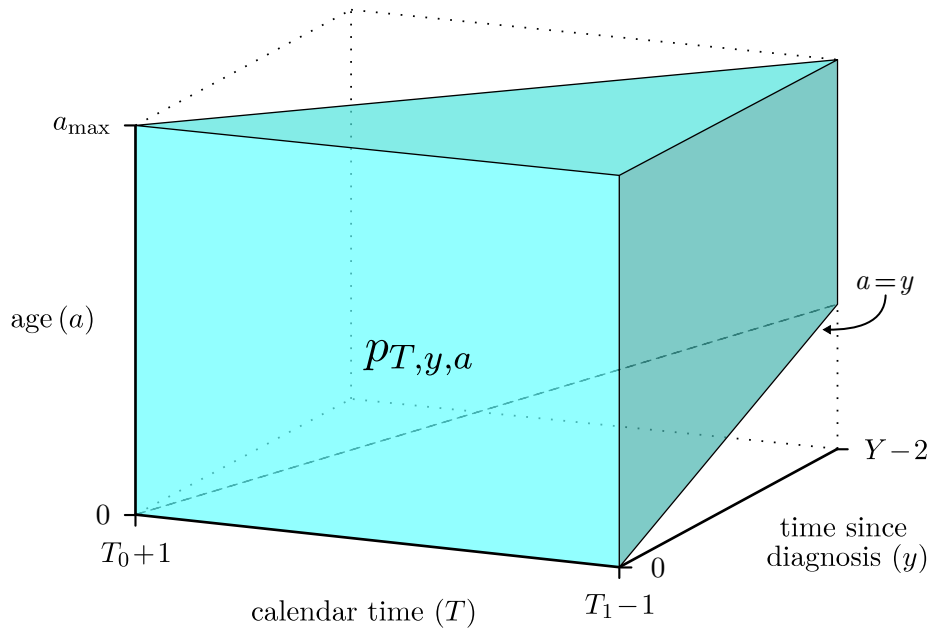


Finally, the extent of the available daily mortality probabilities $p_{T,y,a}$, which are used to estimate the yearly mortality probabilities $M_{T,y,a}$, is very similar to that of the

available prevalence data. The only difference is that there are no data from which to calculate $p_{T_1,y,a}$ and so the restriction on T becomes $T \in [T_0+1, T_1-1]$ and the space of the available daily mortality estimates (Figure 5.4) is given by:

$$\{p_{T,y,a} : T \in [T_0+1, T_1-1], y \in [0, T-T_0-1], a \in [y, a_{\max}]\}\,. \qquad [5.4]$$

Figure 5.4. Three-dimensional extent of available daily mortality data.



5.1.2 Extent of required data

In order to describe the space of required input data, suppose that Y-year limited duration prevalence at the start of year T_1 is known and that the model is to be used to project this to some later time T_2. Then equation [4.6] (restated below as equation [5.5] for convenience) must be evaluated for all $\gamma_2 \in [0, Y-1]$ and $\alpha_2 \in [\gamma_2, a_{\max}]$. By considering this equation, together with equation [4.20] (restated as [5.6]), the spaces of required input incidence, mortality and prevalence can be found.

$$V_{T_2,\gamma_2,\alpha_2} = \begin{cases} N_{T_2,\alpha_2} & \text{if } \gamma_2 = 0 \\ N_{T_2-\gamma_2,\alpha_2-\gamma_2} \prod_{i=0}^{\gamma_2-1} (1 - M_{T_2-\gamma_2+i,i,\alpha_2-\gamma_2+i}) & \text{if } 0 < \gamma_2 < T_2 - T_1 \\ V_{T_1,\gamma_2-T_2+T_1,\alpha_2-T_2+T_1} & \\ \quad \times \prod_{i=\gamma_2-T_2+T_1}^{\gamma_2-1} (1 - M_{T_2-\gamma_2+i,i,\alpha_2-\gamma_2+i}) & \text{if } \gamma_2 \geq T_2 - T_1. \end{cases}$$

[5.5]

$$N_{T,a} = \frac{1}{\tau^2} \sum_{i=1}^{\tau} \Big[\widetilde{N}_{T-1,a}\, i\, (1-\rho_{T-1,a})^{\tau-i+1}\, W_{i,a} + \widetilde{N}_{T-1,a-1}\, (\tau-i)\, (1-\rho_{T-1,a-1})^{\tau-i+1}\, W_{i,a-1} \Big]. \quad [5.6]$$

In equation [5.5], the required known prevalence term is $V_{T,y,a}$, where $T = T_1$, $y = \gamma_2 - T_2 + T_1$ and $a = \alpha_2 - T_2 + T_1$, for $\gamma_2 \geq T_2 - T_1$. Let's assume that $T_2 \leq Y - 1 + T_1$, since otherwise the length of the projection period is such that none of the survivors prevalent at time T_1 can contribute to Y-year limited duration prevalence at time T_2 and the known prevalence term is not required at all.

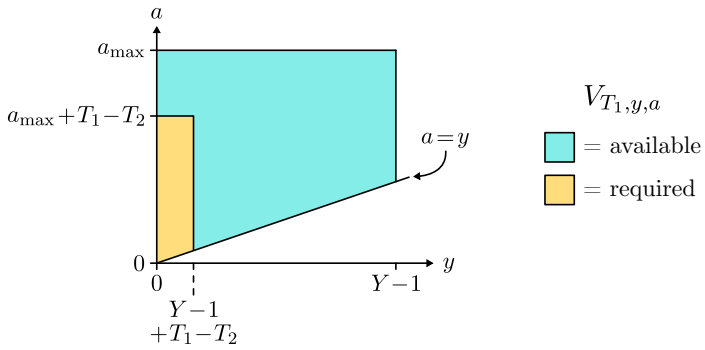
Then the lower bound for the second index of $V_{T,y,a}$ is 0, since $y = \gamma_2 - T_2 + T_1$ and $\gamma_2 \geq T_2 - T_1$; the upper bound is $Y - 1 - T_2 + T_1$, since $\gamma_2 \leq Y - 1$ for Y-year limited duration prevalence. The lower and upper bounds for the third index of $V_{T,y,a}$ are y and $a_{\max} - T_2 + T_1$, respectively, since $a = \alpha_2 - T_2 + T_1$ and $\alpha_2 \in [\gamma_2, a_{\max}]$. Intuitively, the lower bound on the age index a is determined by the fact that nobody may enter the prevalent population in the model before they are born; the upper bound is determined by the fact that those above a certain age at time T_1 will exceed the maximum age $a_{\max}$ at time T_2 and therefore will not be counted as prevalent.

Therefore, the space of required prevalence data is:

$$\{V_{T,y,a} : T = T_1, y \in [0, Y - 1 - T_2 + T_1], a \in [y, a_{\max} - T_2 + T_1]\}. \quad [5.7]$$

This is shown and compared to the space of available prevalence data in Figure 5.5.

Figure 5.5. Comparison of required and available extents of prevalence data.



Inflow terms $N_{T_2-\gamma_2,\alpha_2-\gamma_2}$ are required in equation [5.5] for all $\gamma_2 \in [0, T_2 - T_1 - 1]$ and $\alpha_2 \in [\gamma_2, a_{\max}]$. Therefore, the required space of inflow data is:

$$\{N_{T,a} : T \in [T_1 + 1, T_2], a \in [0, a_{\max} + T - T_2]\}. \quad [5.8]$$

As explained in section 4.4.1, inflow $N_{T,a}$ is the number of people who join the prevalent population at discrete time T and age a, and it can be estimated using continuous time data via equation [5.6]. Consideration of this equation, together with equation [5.8], shows that the required spaces of continuous time incidence counts $\widetilde{N}_{T,a}$, death probabilities $\rho_{T,a}$ and weights $W_{i,a}$ (as defined in equation [4.19]) are given by:

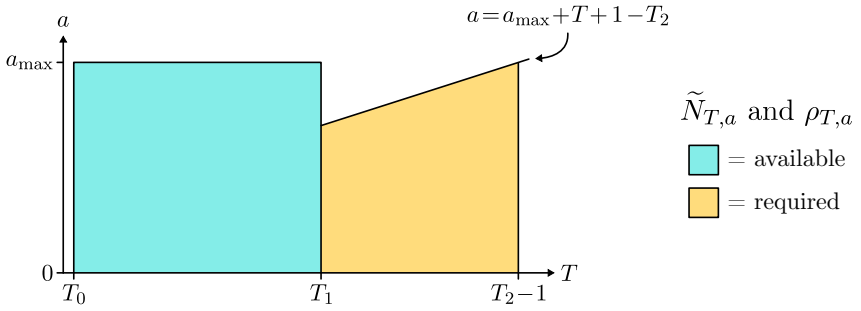
$$\left\{\widetilde{N}_{T,a} : T \in [T_1, T_2 - 1], a \in [0, a_{\max} + T + 1 - T_2]\right\} \quad [5.9]$$

$$\{\rho_{T,a} : T \in [T_1, T_2 - 1], a \in [0, a_{\max} + T + 1 - T_2]\} \quad [5.10]$$

$$\{W_{i,a} : i \in [1, \tau], a \in [0, a_{\max}]\}. \quad [5.11]$$

The spaces of required incidence counts and death probabilities are compared to the available spaces in Figure 5.6.

Figure 5.6. Comparison of required and available extents of continuous time data required to estimate inflow.



The final terms required to evaluate equation [5.5] are the yearly mortality probabilities $M_{T,y,a}$. As before, let's assume that $T_2 \leq Y - 1 + T_1$ so that all lines of equation [5.5] need to be evaluated. Then $M_{T,y,a}$ is required for all $T \in [T_1, T_2 - 1]$, $y \in [0, T + Y - 1 - T_2]$ and $a \in [y, a_{\max} + T - T_2]$. As stated previously, equation [4.18] is used to estimate $M_{T,y,a}$ via $p_{T,y,a}$, the daily mortality probabilities in calendar years subsequent to the year of diagnosis. The above conditions therefore lead to the following space of required daily mortality probabilities:

$$\{p_{T,y,a} : T \in [T_1, T_2 - 1], y \in [0, T + Y - 1 - T_2], a \in [y, a_{\max} + T - T_2]\}. \quad [5.12]$$

Cross-sectional planar and three-dimensional views of this space are shown in Figure 5.7, and compared to the available space of daily mortality data in Figure 5.8.

Figure 5.7. Extent of required daily mortality data.

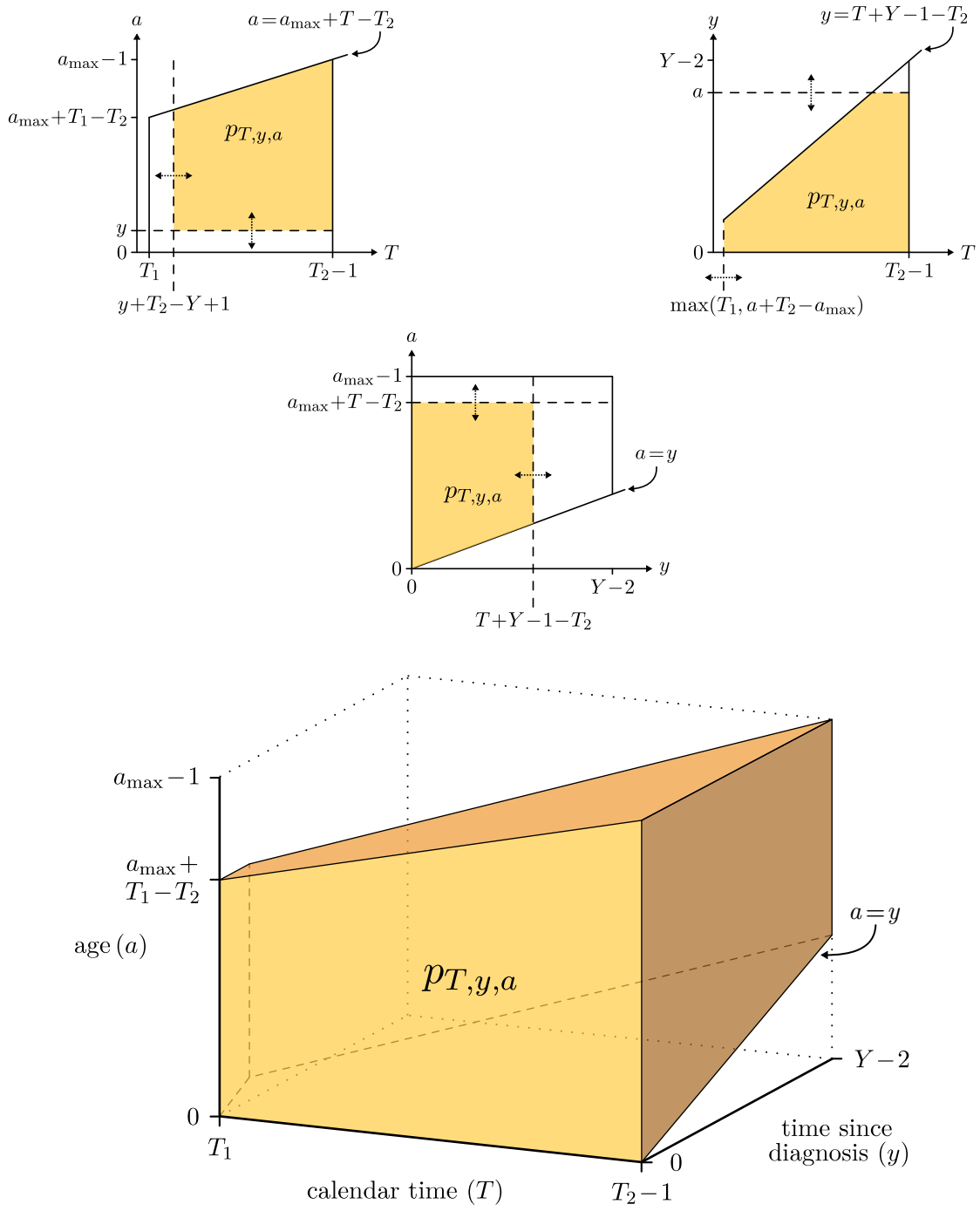
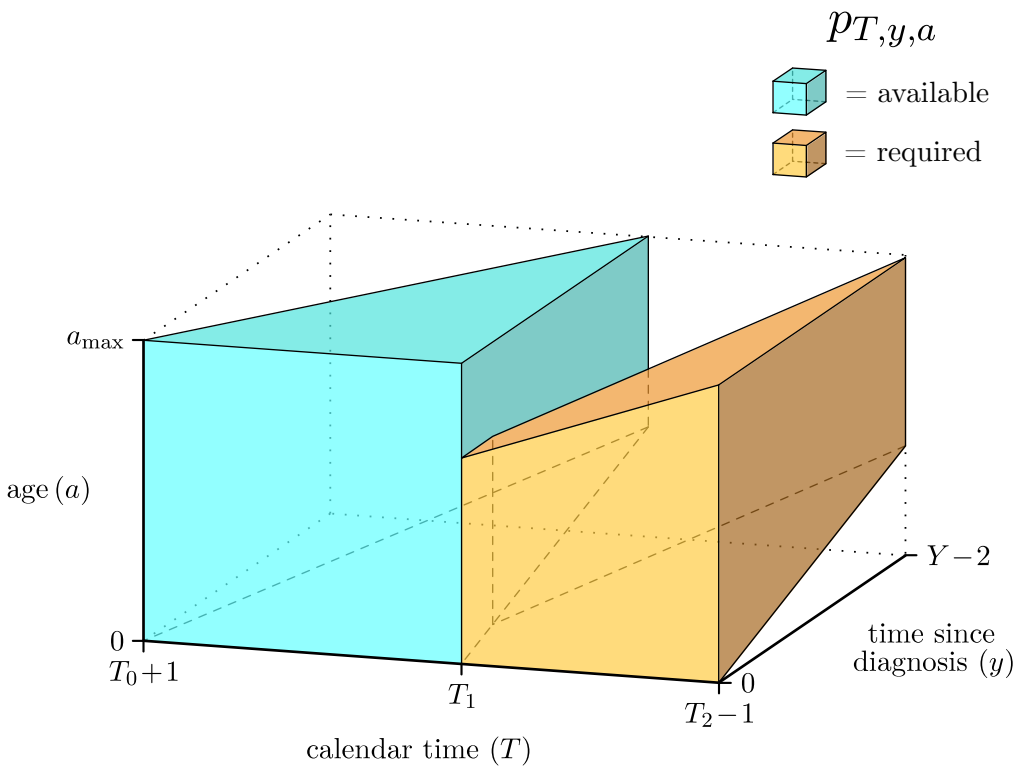


Figure 5.8. Comparison of required and available extents of daily mortality data.



5.2 Available data sources

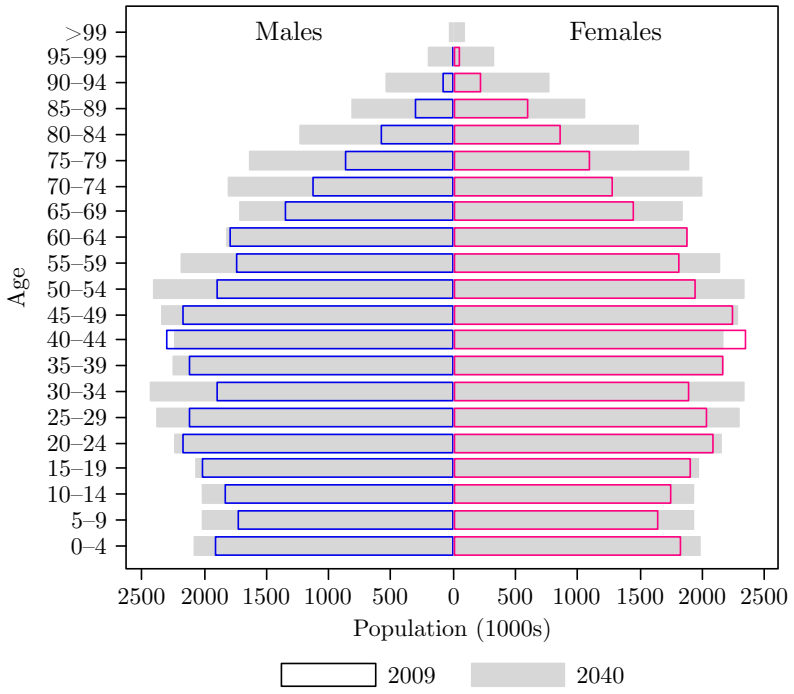
In the previous section the theoretical extent of data available from a generic cancer registry dataset, and the extent of the input data required to make projections of cancer prevalence using the model of Chapter 4, were described. In practice, cancer registry data from the National Cancer Data Repository (National Cancer Intelligence Network, 2011a) were used. This dataset is an amalgamation of data from all eight regional cancer registries in England and as such provides complete geographical coverage of the country. Details of all registered diagnoses of cancer among residents of England in the period 1971 to 2008 (inclusive) were available. In the notation of the previous section, this corresponds to $T_0 = 1971$, $T_1 = 2009$ and $Y = 38$. It is always possible for a small number of cancer registrations never to receive a death notification, leading to so-called 'immortals' in the dataset. The impact of these immortals was minimised by imposing a maximum possible attainable age for all registered cancer patients of 99 years; i.e. in the notation of the previous section, $a_{\max} = 99$.

Consistent with previous chapters, the data were analysed using cohorts of survivors defined by sex and type of cancer – colon, rectum and anus (ICD-10 C18–C21), lung, bronchus and trachea (ICD-10 C33–C34), prostate (ICD-10 C61) and female breast

(ICD-10 C50), as well as all other malignant neoplasms combined (ICD-10 C00–C97 excluding C44 and those codes mentioned previously).

Historical and estimated future national population data were supplied by ONS. The 'principal' (i.e. the most likely) 2008-based projections of the size of the population of England and the UK, by age, sex and year up to 2040, were used (Office for National Statistics, 2011b). A population pyramid showing the size and age structure of the population of England in 2009 and the projected population for 2040 is shown in Figure 5.9. It can be seen that the population is anticipated to increase in almost all age groups, for both sexes, with the largest increases in the oldest age groups; i.e. the population is both growing and ageing. Specifically, the size of the population of the UK is anticipated to grow by almost 20% from 62 million in 2009 to 74 million in 2040, and the median age to increase from approximately 37 to 41 years (males) and from 40 to 43 years (females). The proportion of the population at least 65 years of age is expected to increase from 16% in 2009 to 24% in 2040.

Figure 5.9. Population pyramid for the UK, 2009 and 2040.



Source: Office for National Statistics

5.3 Regression models

All required prevalence input data are known if it is assumed that the cancer registry data contain Y years of cancer registrations prior to year T_1 and that the task is to

project Y-year limited duration prevalence from the start of year T_1 to the start of year T_2. Future incidence rates and mortality probabilities are not known but can be estimated from the available data by extrapolating from existing empirical trends. Regression models were designed for this purpose.

Using the cancer registry dataset described above, 38-year prevalence at the start of 2009 was known, and this was to be projected to the year 2040; i.e. $T_1 = 2009$, $Y = 38$ and $T_2 = 2040$. With these parameters, the extents of the available and required input data can be seen in the figures of section 5.1.2 and these visualisations informed the design of the regression models.

The following models were chosen for their simplicity and general applicability to all cancer type and sex combinations. Similar models were chosen for both incidence and mortality and, in all cases, age groups under 30 years were excluded from the regression due to problems surrounding small numbers – incidence rates and mortality probabilities were assumed to be constant in these age groups from 2009 to 2040. This assumption was considered to be reasonable given the types of cancer being studied (prostate, breast, colorectal and lung) for which the majority of disease is diagnosed in the older age groups. The 'all other' category does contain a variety of cancer types for which this assumption would not be appropriate (bone cancers, lymphoma, leukaemia, etc.), but since these are not being studied individually the effect of the assumption in this category is minimal.

The regression models were initially evaluated on a sub-set of the available cancer registry data as part of an exercise to test the prevalence projection model; this is described in section 5.4. Only after this exercise was considered to be complete was the whole dataset used, and at this point it was necessary to make an alteration to the incidence regression model, as described in section 5.5.

5.3.1 Incidence

Incidence rates were projected in 5-year age groups and 5-year calendar time periods using an age-period-cohort Poisson regression model with a log link function. This is a log-linear regression model commonly used to project incidence counts and rates, as well as other epidemiological statistics (Clayton and Schifflers, 1987). The number of incident cases of disease is assumed to follow the Poisson distribution and the logarithm of its expected value is modelled as a linear combination of the explanatory variables age, period and cohort. If a rate is being modelled then the logarithm of the population

at risk is also included as an explanatory variable in the regression equation with a fixed coefficient of 1 – this is known as the 'offset' variable.

Age-period-cohort models can be constructed in a variety of ways – for example, the cohort or period term may be omitted, and a linear 'drift' term may or may not be included (Clayton and Schifflers, 1987). A drift term is used to model the general trend in the response variable over time that applies equally in the period and cohort dimensions – it may therefore be specified as either a period-drift or a cohort-drift. In many instances, different constructions result in different parameterisations of the exact same model; this is known as the identifiability problem and is caused by the fact that age, period and cohort are not linearly independent variables. When modelling cancer incidence rates, the cohort term is often preferred to the period term since long-term habits and historical exposures to risk factors are likely to vary by birth cohort rather than time period (Robertson et al., 1999). For this reason, age group and birth cohort (defined as period minus age group) were used as categorical explanatory variables. To allow projections of incidence rates beyond the most recent year of available data in all age groups, period was also included as a linear explanatory variable (the period-drift term). The coefficients of the regression equation were estimated using the `proc genmod` procedure in the SAS statistical programming package (SAS Institute Inc., Cary, NC, USA).

The variables used in the incidence regression were at a resolution of five years in order to minimise any issues with small numbers that may have caused unstable projections. The model for projecting cancer prevalence, however, operates at a 1-year resolution and so linear interpolation was used to convert the projected incidence rates in 5-year periods and age groups to incidence rates in 1-year periods and age groups. Projected populations for England, as supplied by ONS, were then used to convert the projected incidence rates to numbers for use in the prevalence projection model.

5.3.2 Mortality

The daily probability of death among survivors of a given age and time since diagnosis in a given time period was estimated using a person-time analysis of the cancer registry data which calculated the total number of deaths and divided this by the total number of person-days at risk. Two separate Poisson regression models were then used to project the daily mortality probabilities $\rho_{T,a}$ and $p_{T,y,a}$ into the required spaces defined by equations [5.10] and [5.12], respectively. In both cases, the response variable was the number of deaths and the logarithm of the number of person-days at risk was the offset

variable. An arbitrary number of person-days at risk was used where no empirical data were available (i.e. in the projected period) – this was appropriate since only the probability of death is required to be projected, not the actual number of deaths. The choice of explanatory variables was different in each of the two models, as explained below.

The daily mortality probability in the calendar year of diagnosis, $\rho_{T,a}$, was projected using an age-period model; 10-year age group and 5-year period were the explanatory variables, with the latter being a linear drift term as in the incidence model. The weights $W_{i,a}$ that are required to scale $\rho_{T,a}$ (as per equation [4.19]) were estimated for each cancer type and sex combination using the most recent 10 years of cancer registry data, i.e. diagnoses made between 1999 and 2008 (inclusive), under the assumption that the weights themselves remain constant over calendar time. They were also assumed to be constant within broad age groups (0–44, 45–64 and ≥65 years) and within each month post-diagnosis.

The daily mortality probabilities in years subsequent to the year of diagnosis, $p_{T,y,a}$, were projected using 10-year age group, 5-year period (again as a linear drift) and 1-year time since diagnosis as the explanatory variables. In addition, an age–time since diagnosis interaction term was included to account for the fact that, as time since diagnosis increases beyond the initial phase of the disease, the effect of general background mortality becomes more pronounced in the older age groups than in the younger age groups – i.e. the risk of dying from a non-cancer cause increases with age. The interaction groups were defined following an empirical investigation; time since diagnosis was grouped in one year intervals up to 5 years, then 5–9 years and ≥10 years, and age was grouped as <60, 60–69, 70–79 and ≥80 years. Once again, the coefficients of the regression equation were estimated using the `proc genmod` procedure in the SAS statistical programming package (SAS Institute Inc., Cary, NC, USA).

Mortality data were grouped at various resolutions for the purposes of regression analysis (1-year time since diagnosis bands, 5-year calendar time periods and 10-year age groups), and therefore some interpolation was required to transform the output to the 1-year resolution required by the prevalence projection model. As with the incidence data, linear interpolation was used for this task.

5.4 Evaluation exercises

Exercises were undertaken to test the prevalence projection model and to explore its forecasting ability. During this evaluation stage, it was desirable to maintain the integrity of both the model and the data, so that no changes were made based on specific features of the data that would result in over-fitting of the model. A sub-set of the cancer registry data for England was therefore created for testing purposes only. This consisted of a 50% random sample of all male colon cancers (ICD-10 C18) diagnosed in the period 1972–2006 (inclusive), approximately 122,000 records.

The first exercise tested the construction of the prevalence projection model, and its implementation using cancer registry data as input. Using the test data, 20-year prevalence at the start of 1997 was estimated and the model was used to project this 10 years forward to 2007 using empirical incidence and mortality data from the period 1997–2006. Since all required input data was available empirically, and the intent was to test the construction of the projection model, no regression was carried out. An empirical estimate of 20-year prevalence at the start of 2007 was then calculated using the test data, and compared with the results of the model projections.

No discrepancies were found that indicated a structural or systematic error in the VBA code used to implement the model, and the two estimates of prevalence were in very close agreement. Table 5.1 shows the absolute and percentage differences between the projected and empirical estimates of the number of survivors. The largest percentage differences were observed in the youngest age group (0–44 years), due largely to the small numbers involved, but, nonetheless, overall the difference was very small (1.94%). In the other age groups, the differences were negligible: 0.28% overall in the 45–64 years age group and 0.32% overall in the ≥65 years age group.

Table 5.1. Results of exercise 1. Absolute and percentage differences between projected and empirical number of survivors, by attained age and time since diagnosis.

	Age group							
	0–44		45–64		≥65		All ages	
Years since diagnosis	Absolute	%	Absolute	%	Absolute	%	Absolute	%
<1	1	0.94	0	0.00	31	1.15	32	0.85
1–5	-1	-0.46	9	0.44	28	0.41	36	0.40
5–20	10	5.21	5	0.26	4	0.04	19	0.15
All (0–20)	10	1.94	14	0.28	63	0.32	87	0.35

Projecting prevalence 10 years forward in time using true incidence and mortality data for the projected period resulted in almost exactly the same prevalence estimates as were obtained by a count of empirical data. The results of this exercise therefore showed that the discrete time mechanics of the model for projecting cancer prevalence, and its handling of real world continuous time data input, worked well.

Having established this, the exercise was then repeated using projected incidence and mortality for the period 1997–2006, rather than the empirical data. Incidence and mortality in the period 1972–1996 were projected forward to 2006 using the regression models described in section 5.3. The absolute and percentage differences between the projected and empirical number of survivors in 2007 using these input data are shown in Table 5.2.

Table 5.2. Results of exercise 2. Absolute and percentage differences between projected and empirical number of survivors, by attained age and time since diagnosis.

	Age group							
	0–44		45–64		≥65		All ages	
Years since diagnosis	Absolute	%	Absolute	%	Absolute	%	Absolute	%
<1	-21	-19.81	-18	-1.92	381	14.07	342	9.12
1–5	-37	-17.05	-120	-5.81	611	9.03	454	5.02
5–20	-14	-7.29	28	1.43	287	2.82	301	2.45
All (0–20)	-72	-13.98	-110	-2.22	1,279	6.51	1,097	4.37

The differences are now much greater than when using only empirical input data. This was to be expected, since the regression models provide only an estimate of incidence and mortality in the period 1997–2006 based on the trends observed in the period 1972–1996. However, this exercise does give a measure of the predictive power of the model for this particular test set of data when using the regression procedures described.

Once again, it was in the youngest age group (0–44 years) that the percentage differences between the projected and empirical estimates of prevalence were greatest; overall the projected number was 14% lower than the empirical. This is partly due to the small numbers involved in this age group, since the instability in the empirical incidence and mortality cannot be captured by the regression models (which produce smooth estimates). It is also a direct result of the assumption that incidence rates and mortality probabilities remained constant in the very youngest age groups, the under 30s.

Of note is the apparent age structure to the percentage differences in Table 5.2 – prevalence in the youngest age groups was underestimated, and in the oldest age groups it was overestimated. This was shown to be not a structural problem with the prevalence model, but a result of the success or failure of each regression procedure (or assumption of constancy for the under 30s) to correctly predict the input data for the period 1997–2006. Substituting the projected incidence data for the empirical incidence data, and running the model, removed the underestimation of prevalence in the youngest age groups, but the overestimation in the oldest age groups remained. Conversely, running the model with projected incidence and empirical mortality removed the overestimation of prevalence in the oldest age groups whilst leaving the underestimation in the youngest age groups. The mortality regression model had underestimated the death rates in the oldest age groups, particularly those over 80 years, and this is what caused the increasing overestimation of prevalence as age increased.

It was therefore shown that, unsurprisingly, the accuracy of the model's projections of future cancer prevalence was dependent upon the performance of the regression models used to estimate the input data (future incidence and mortality). However, this performance can only be assessed retrospectively with the type of exercises and historical data described above. It is, of course, impossible to predict exactly what future incidence and mortality will be, but the regression models used here resulted in a projected total number of survivors (all age and time since diagnosis groups combined) that was within 5% of the empirical number (Table 5.2).

5.5 Incidence and mortality projections 2009–2039

Having completed the testing and evaluation exercises, the input data required to project cancer prevalence in England from 2009 to 2040 was prepared. The full extent of the available cancer registry data (as described in section 5.2) was analysed, and incidence and mortality in the period 1974–2008 was projected to 2039 using the regression models described in section 5.3. Data pertaining to the period 1971–1973 were excluded from the regression since they did not cover a complete five year period.

The incidence regression model was, however, found to be inappropriate for such long-term projections. During the testing and evaluation exercises, incidence was projected no more than 10 years forward, but for the main analysis 30 year projections were required. It was found that the inclusion of the categorical cohort term resulted in unstable projections because for many cancer types the regression procedure resulted in

cohort effects acting in one direction and a period drift effect acting in the opposite direction – for example, for male lung cancer the best fit to the data was obtained with all cohort coefficients being negative, but the period-drift coefficient being positive. This is problematic for long-term projections since coefficients for cohorts not present in the empirical data are not estimable and must therefore be ignored or remain constant based on the nearest estimable cohort parameter; this can result in an undesirable false reversal of trends many years into the future, based not on any empirical evidence, but on a failure of the regression model. For this reason, the incidence regression model was simplified and the cohort term was removed so that all cohort and period effects were condensed and estimated by a single period-drift coefficient.

5.5.1 Incidence results

Figure 5.10 shows the results of the incidence regression described above, for each cancer type and sex combination. The outlines of graphs for age groups which (individually) contained less than 10% of the total number of registered cancer diagnoses in the period 2004–2008 for each cancer type and sex combination have been greyed out – this is to provide an indication of the performance of the regression model in the age groups for which the inflow to the prevalent population is likely to be greatest.

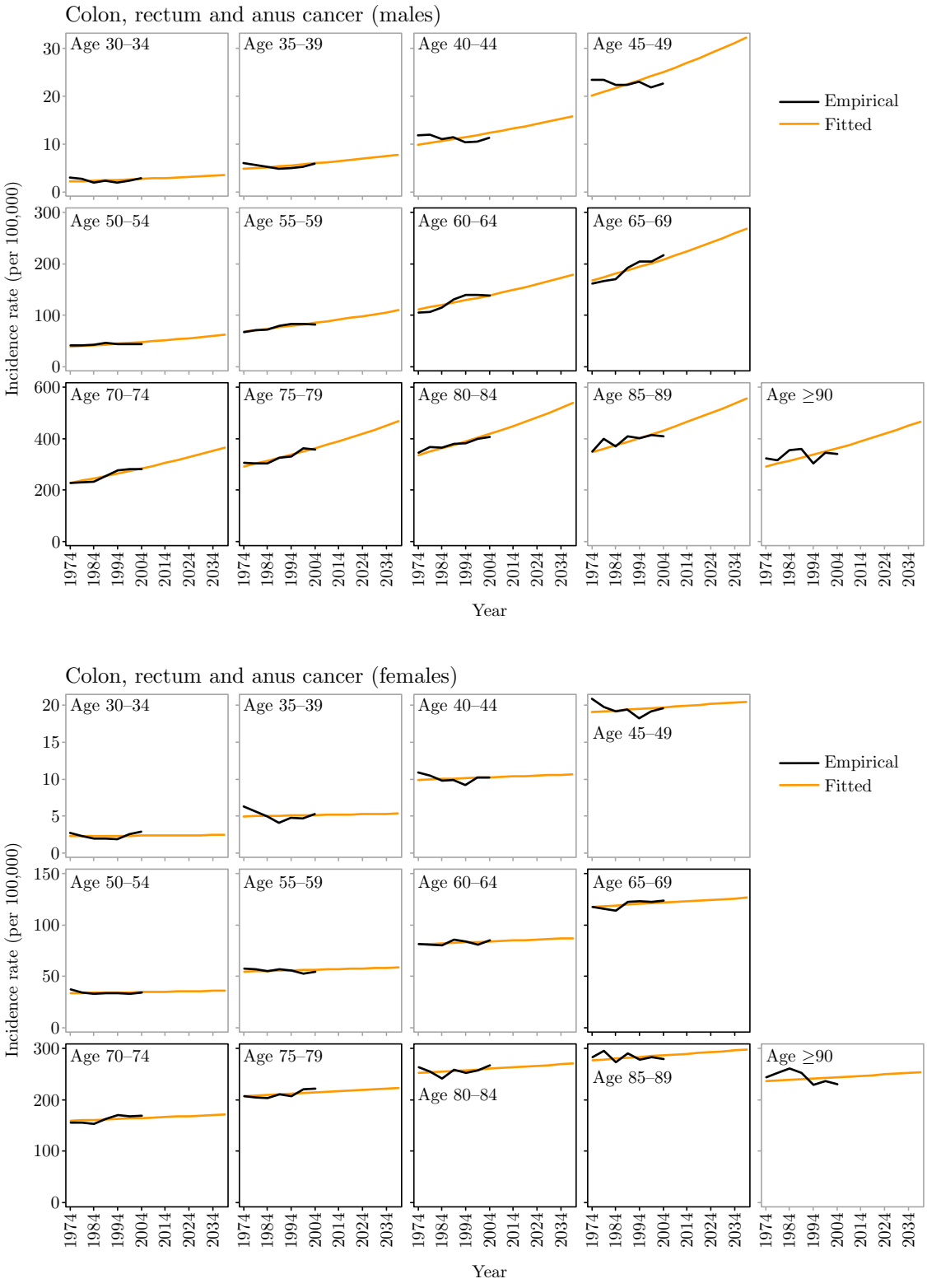
The majority of incidence rates increased in the period 1974–2008 and, accordingly, the regression model extended these increasing trends to 2039. The notable exceptions were male lung cancer which showed a decreasing trend – primarily as a result of dramatically decreasing smoking prevalence among men in the UK since the 1970s (Davy, 2006) – and female colorectal cancer which had largely stable incidence rates in most age groups.

The largest increases in incidence rates were seen for prostate cancer. For example, in the age group 70–74 years, incidence rates rose from 200 cases per 100,000 population in the period 1974–1978 to 625 per 100,000 in 2004–2008; a more than three-fold increase in 30 years. Naively extending this trend with the log-linear period-drift resulted in an estimated incidence rate of almost 3,000 per 100,000 by 2039. This is certainly an overestimate of the likely future incidence rates of prostate cancer. Much of the historical increase in prostate cancer incidence is attributable to the introduction in the early 1990s of the PSA test as a screening tool for prostate cancer. This test essentially re-framed the definition of the disease, but its effect on incidence rates cannot realistically be expected to continue in the same way until 2039. This is a

limitation of the described approach to projecting incidence rates – it simply extends existing trends into the future and in some cases this may be clearly unrealistic. Other methods for projecting incidence rates are possible – for example, by including explicit assumptions regarding the effects of screening programmes, as in Mistry et al. (2011) – but for the purposes of this work a model was sought that could easily and reasonably be applied to all cancer types. Additionally, instead of assuming existing trends will continue unabated, other authors have advocated applying arbitrary or empirically motivated attenuation factors to recent trends in order to project cancer incidence into the future (Møller et al., 2002; Mistry et al., 2011). This approach was not used in this thesis, however, since alternative scenarios for future incidence (and mortality), and their effect on projections of prevalence, are considered separately in Chapter 6.

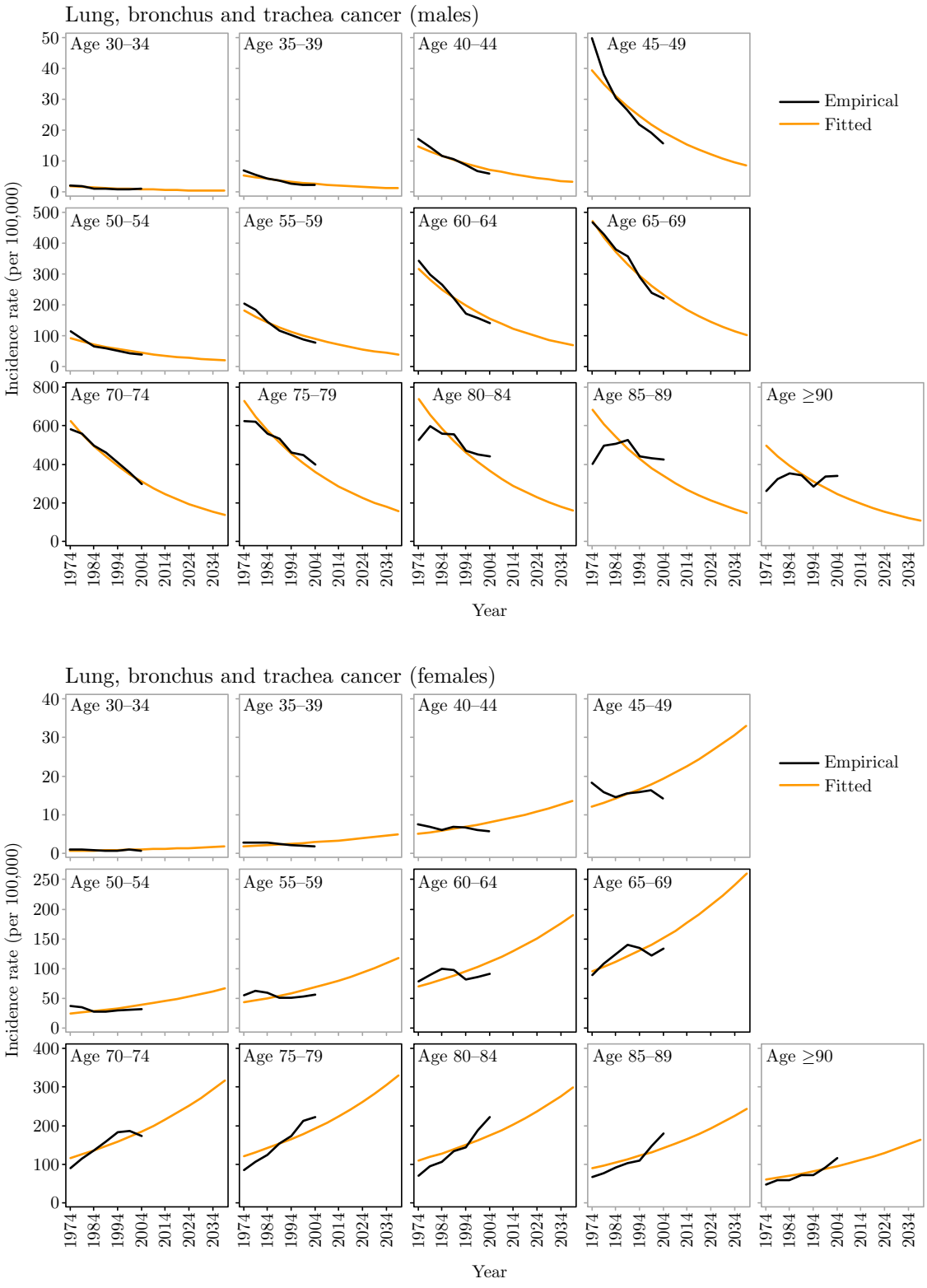
The residuals of the regression procedure (not displayed here) were analysed to assess the goodness of fit of the model to the empirical data. For all cancer type and sex combinations the standardised residuals were quite well normally distributed and showed little correlation with the explanatory variables. Residuals were largest for incidence of lung cancer. The female lung cancer results illustrate a limitation imposed by the simplicity of the regression model. The empirical data show that incidence rates generally increased in the period 1974–2008 for all age groups older than 65 years, were quite steady in the age groups 50–64, and decreased slightly in the age groups under 45. However, the regression model is unable to account for such different period trends in each age group since there is only one period-drift parameter and no period-age interaction term. This results in female lung cancer incidence rates being projected to increase in all age groups, even though the empirical trends do not support this. The effect of this limitation on projections of overall cancer prevalence is likely to be small, however, since a large majority of cancers are diagnosed in older age groups for which the regression model performed satisfactorily. This problem is much less significant for the other types of cancer, and is accepted as a limitation brought about by the desire to use the same regression model for every cancer type and sex combination. Nonetheless, an area of further study would be to develop individual regression models for each sex and type of cancer.

Figure 5.10. Empirical and projected incidence rates, England, 1974–2039; by cancer type, sex and age group*.



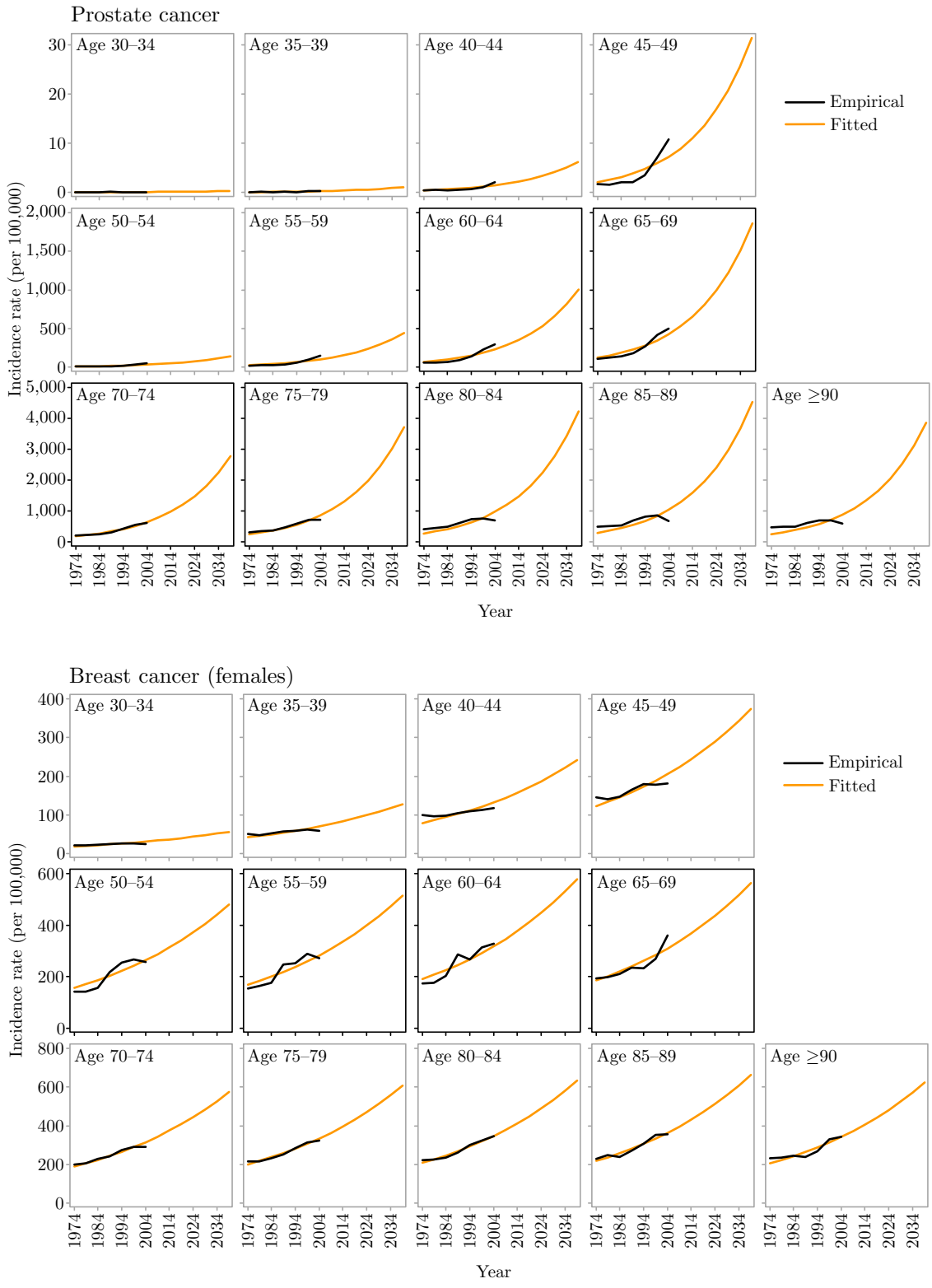
*Outlines of graphs for age groups that contain less than 10% of the total number of cancer diagnoses in the period 2004–2008 have been greyed out.

Figure 5.10 (continued). Empirical and projected incidence rates, England, 1974–2039; by cancer type, sex and age group*.



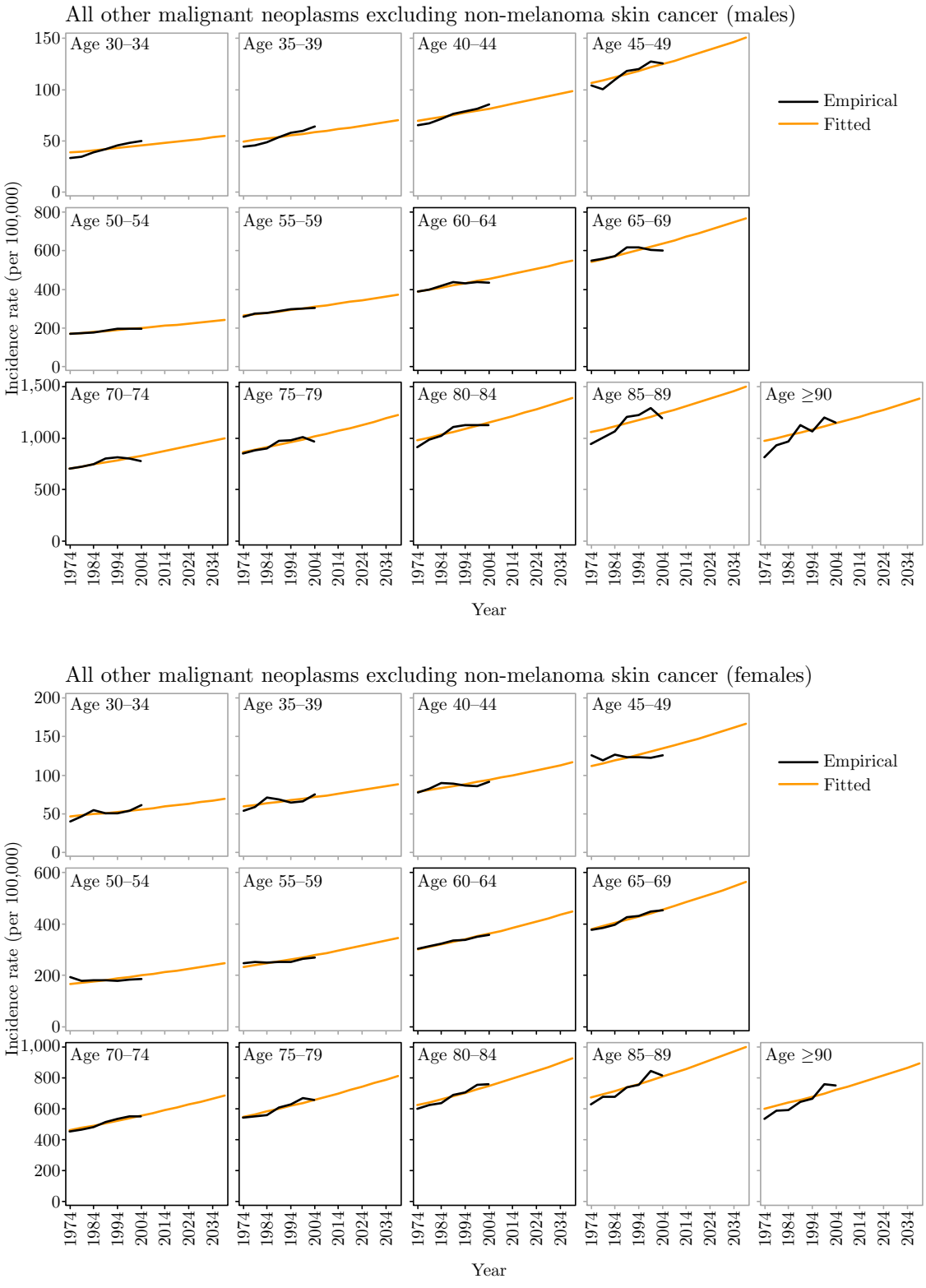
*Outlines of graphs for age groups that contain less than 10% of the total number of cancer diagnoses in the period 2004–2008 have been greyed out.

Figure 5.10 (continued). Empirical and projected incidence rates, England, 1974–2039; by cancer type, sex and age group*.



*Outlines of graphs for age groups that contain less than 10% of the total number of cancer diagnoses in the period 2004–2008 have been greyed out.

Figure 5.10 (continued). Empirical and projected incidence rates, England, 1974–2039; by cancer type, sex and age group*.



*Outlines of graphs for age groups that contain less than 10% of the total number of cancer diagnoses in the period 2004–2008 have been greyed out.

5.5.2 Mortality results

Figure 5.11 shows the empirical and projected average daily probability of death in the calendar year of diagnosis for cancer survivors (i.e. $\rho_{T,a}$), by 10-year age group and 5-year period. Similarly to Figure 5.10, the outlines of graphs for age groups which contained the smallest numbers of registered cancer diagnoses have been greyed out – this time, since the age bands are wider, the criterion for being greyed out is less than 20% of the total number of registered cancer diagnoses in the period 2004–2008.

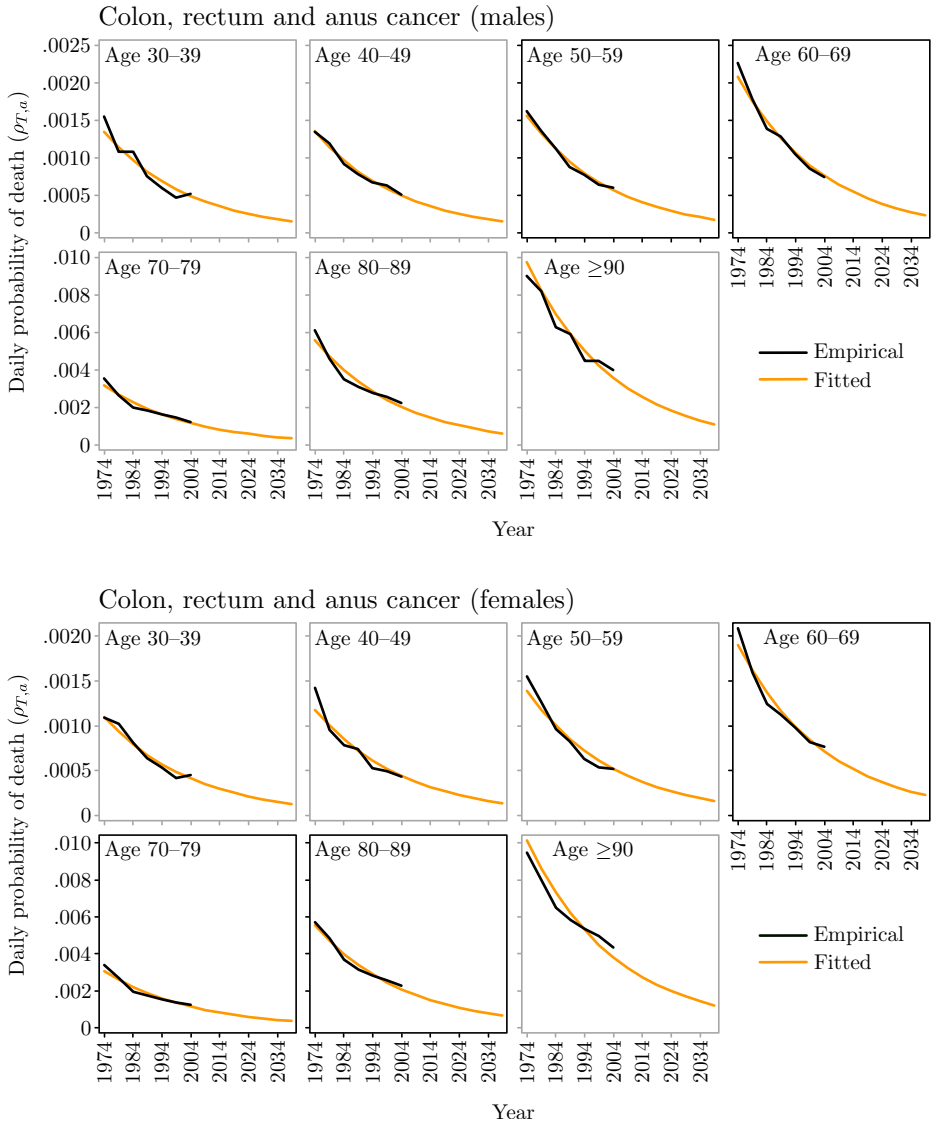
In all age groups, for every cancer type and sex combination, $\rho_{T,a}$ showed a decreasing trend. This reflects the great improvements in cancer treatment that have been made in the last 40 years, as well as the effects of earlier diagnosis, both of which have contributed to increasing cancer survival particularly in the first year following diagnosis (Rachet et al., 2009). The fit of the regression model to the data was observed to be reasonably good in most instances, with small residuals.

Figure 5.12 shows the results of the second mortality regression model which projected the average daily probability of death in calendar years subsequent to the year of diagnosis (i.e. $p_{T,y,a}$), in 10-year age groups, 5-year time periods and 1-year time since diagnosis bands. Plotted are the yearly mortality probabilities, $M_{T,y,a}$, calculated from the daily probabilities using equation [4.18]. In order to maintain the readability of the graphs, only values for $y = 0$ (solid lines) and $y = 5$ (dashed lines) are displayed. As in Figure 5.11, the outlines of graphs for age groups which contained less than 20% of the total number of registered cancer diagnoses in the period 2004–2008, for each cancer type and sex, have been greyed out.

As with the mortality probabilities in the calendar year of diagnosis, these probabilities decreased over time for all cancer type, sex and time since diagnosis combinations, reflecting the general lengthening of cancer survival time in England since the 1970s. For the majority of cancers, the fit of the regression model to the observed data was acceptable and the projections plausible. Residuals were larger than those produced by the incidence regression, mainly due to the necessary inclusion of the time since diagnosis term (y) at a 1-year resolution which resulted in rather small numbers and unstable estimates; for example, for prostate cancer in the relatively young age groups. In the period 1974–2008, mortality probabilities for breast cancer survivors decreased at a much faster rate in the younger age groups compared with the older age groups, and this was a characteristic that was particularly hard for the regression model to account

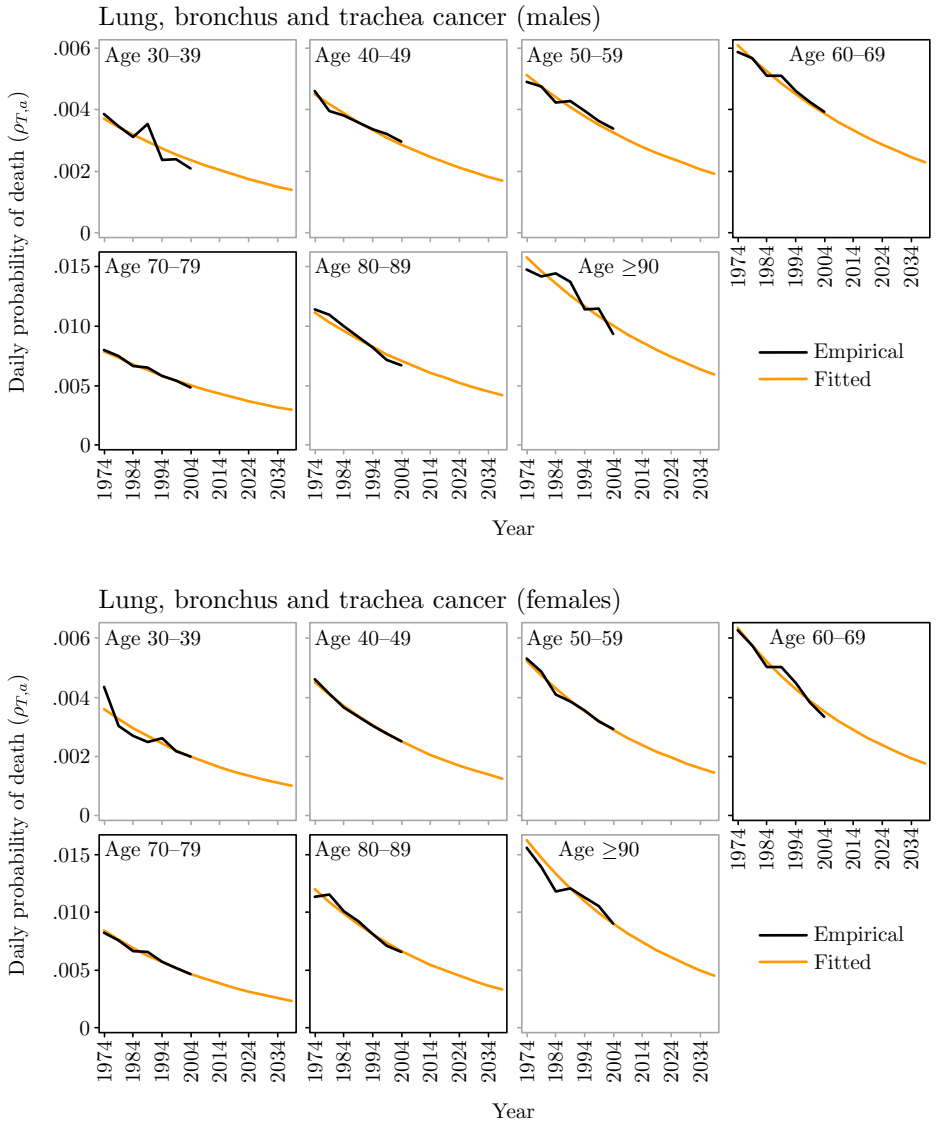
for. The inclusion of the age–time since diagnosis interaction term did, however, improve the fit to the empirical data in this case.

Figure 5.11. Empirical and projected average death probabilities in the calendar year of diagnosis for cancer survivors, England, 1974–2039; by cancer type, sex, age group*.



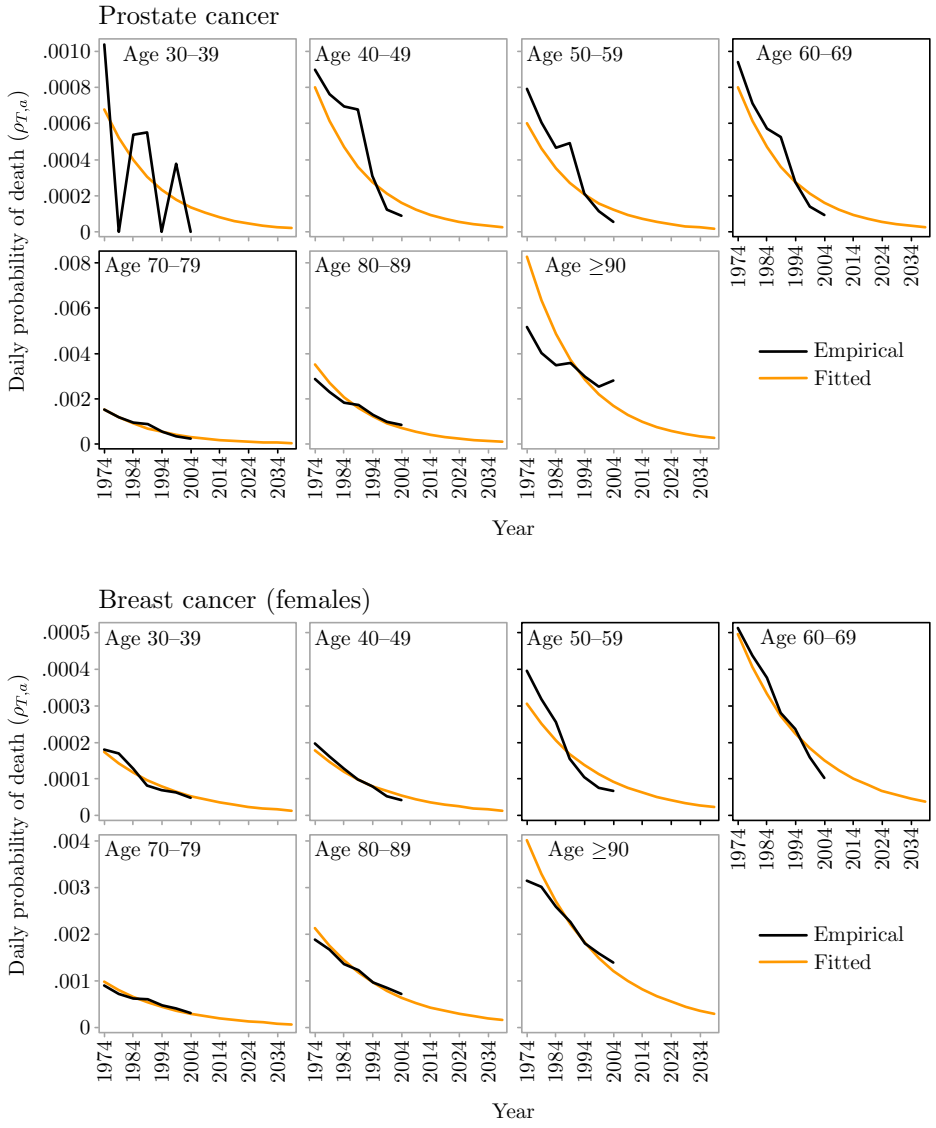
*Outlines of graphs for age groups that contain less than 20% of the total number of cancer diagnoses in the period 2004–2008 have been greyed out.

Figure 5.11 (continued). Empirical and projected average death probabilities in the calendar year of diagnosis for cancer survivors, England, 1974–2039; by cancer type, sex, age group*.



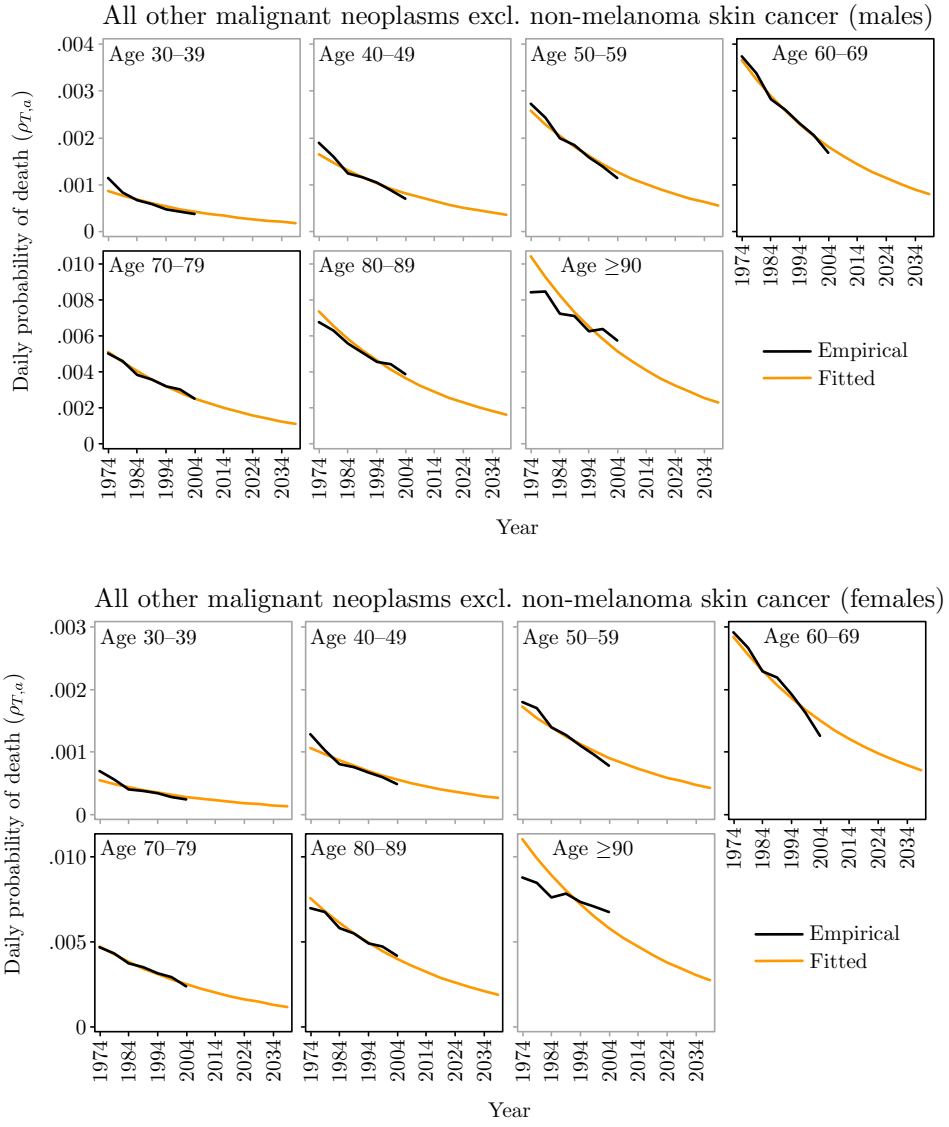
*Outlines of graphs for age groups that contain less than 20% of the total number of cancer diagnoses in the period 2004–2008 have been greyed out.

Figure 5.11 (continued). Empirical and projected average death probabilities in the calendar year of diagnosis for cancer survivors, England, 1974–2039; by cancer type, sex, age group*.



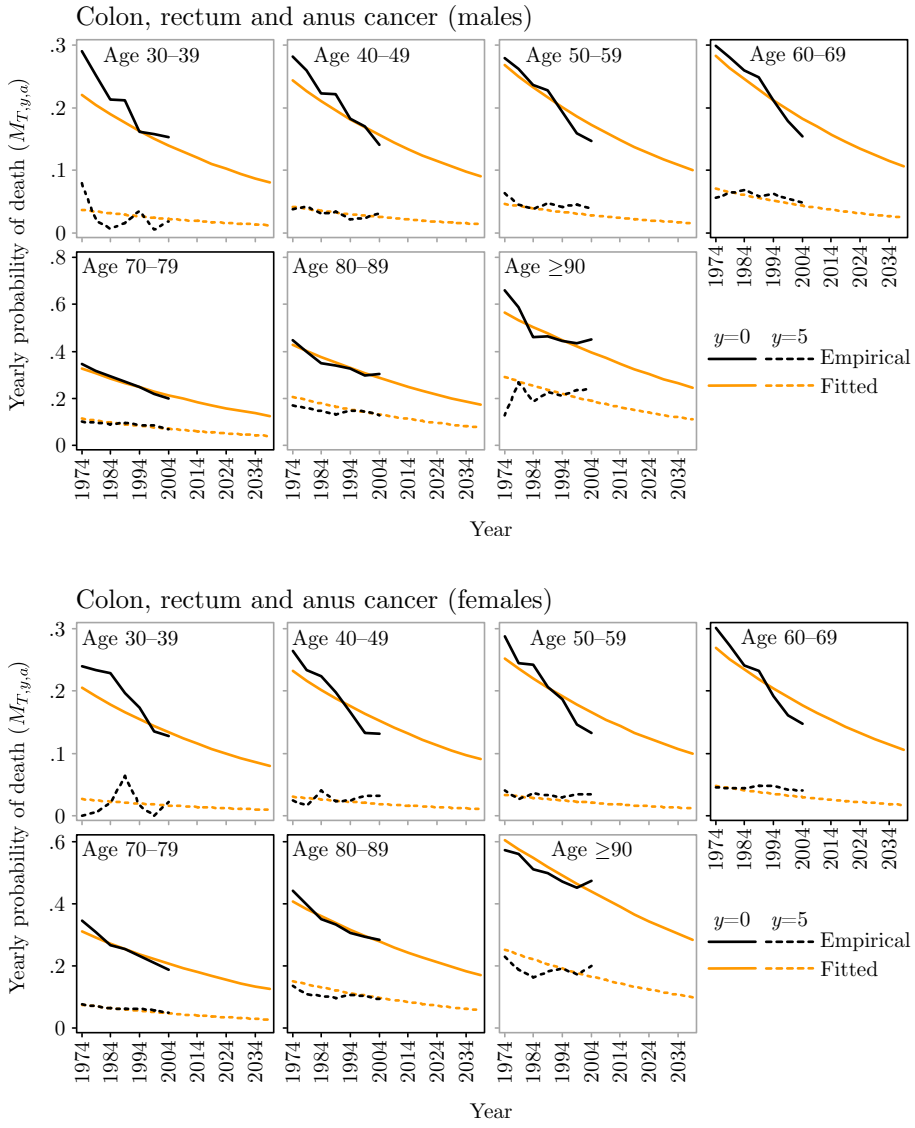
*Outlines of graphs for age groups that contain less than 20% of the total number of cancer diagnoses in the period 2004–2008 have been greyed out.

Figure 5.11 (continued). Empirical and projected average death probabilities in the calendar year of diagnosis for cancer survivors, England, 1974–2039; by cancer type, sex, age group*.



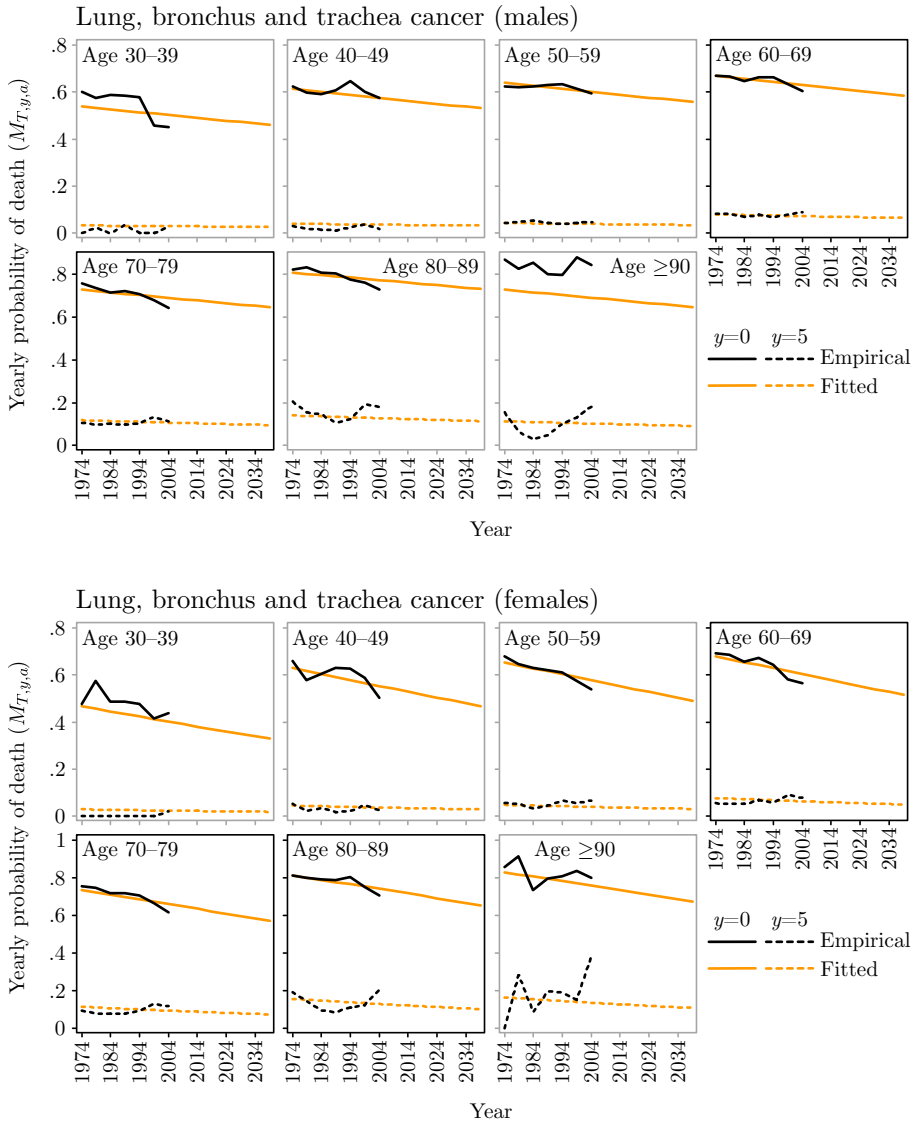
*Outlines of graphs for age groups that contain less than 20% of the total number of cancer diagnoses in the period 2004–2008 have been greyed out.

Figure 5.12. Empirical and projected average yearly death probabilities for cancer survivors, England, 1974–2039; by cancer type, sex, age group* and time since entry to the prevalent population.



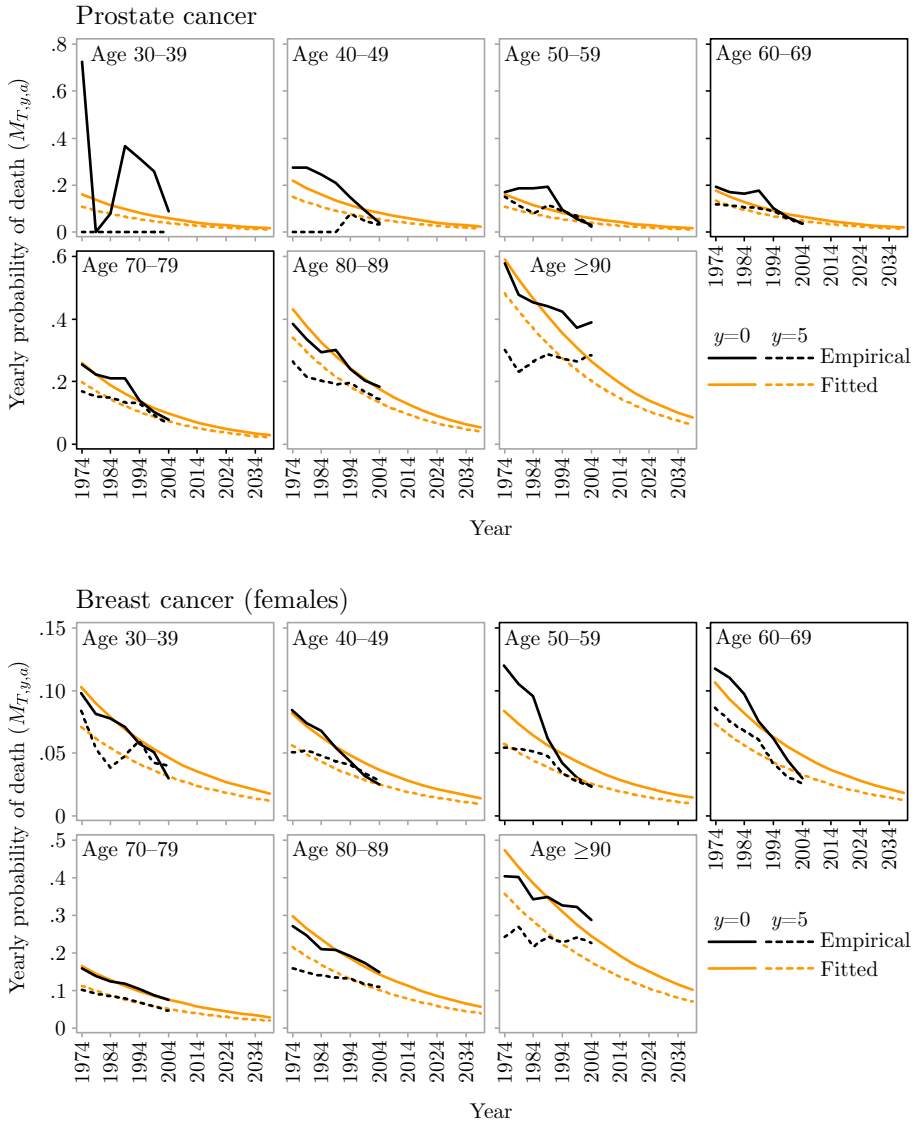
*Outlines of graphs for age groups that contain less than 20% of the total number of cancer diagnoses in the period 2004–2008 have been greyed out.

Figure 5.12 (continued). Empirical and projected average yearly death probabilities for cancer survivors, England, 1974–2039; by cancer type, sex, age group* and time since entry to the prevalent population



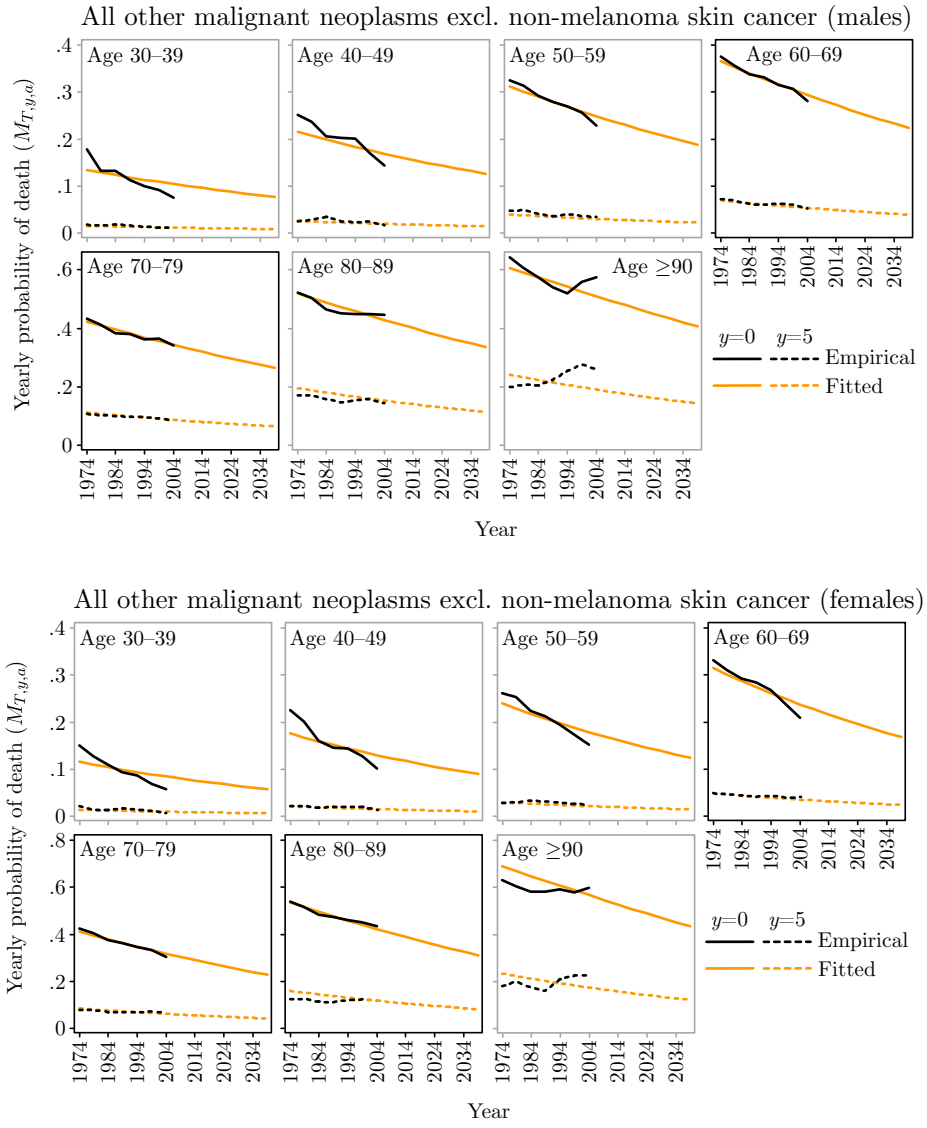
*Outlines of graphs for age groups that contain less than 20% of the total number of cancer diagnoses in the period 2004–2008 have been greyed out.

Figure 5.12 (continued). Empirical and projected average yearly death probabilities for cancer survivors, England, 1974–2039; by cancer type, sex, age group* and time since entry to the prevalent population



*Outlines of graphs for age groups that contain less than 20% of the total number of cancer diagnoses in the period 2004–2008 have been greyed out.

Figure 5.12 (continued). Empirical and projected average yearly death probabilities for cancer survivors, England, 1974–2039; by cancer type, sex, age group* and time since entry to the prevalent population



*Outlines of graphs for age groups that contain less than 20% of the total number of cancer diagnoses in the period 2004–2008 have been greyed out.

5.5.3 Discussion

Regression models were chosen that would be generally applicable to all cancer types so that the production of prevalence projections was not unnecessarily complex. The fit of the regression models to the observed data could be improved; for example, by tailoring the regression procedures based on characteristics and trends observed in the data particular to each cancer type. However, in most cases, the fit to the observed data was acceptable and the projected incidence and mortality estimates were plausible. Indeed, for the purposes of projecting cancer prevalence, it is more important to provide roughly plausible estimates of future incidence and mortality than it is to have a good fit to the observed data, since it is only the future estimates that are required by the prevalence projection model (see Figures 5.6 and 5.8).

5.6 Summary

In this chapter the extent of input data necessary for projecting cancer prevalence was explored. Regression models were developed to provide the required estimates of future incidence and mortality based on empirical trends. Using a test dataset, evaluation exercises were conducted and revealed that the mechanics of the discrete time prevalence model worked properly, dealt with continuous time data inputs appropriately and were coded in the Excel VBA computer program correctly.

National cancer registry data for England (covering all diagnoses made in the period 1971–2008) were used to generate input data that would allow projections of cancer prevalence up to the year 2040. The results of these projections are detailed and discussed in the next chapter. Alternative scenarios of future incidence and mortality (i.e. not based on existing trends), and their effect on future cancer prevalence, are also considered.

Chapter 6. Projections of cancer prevalence to 2040

Using the model and input data described in the previous two chapters, projections of cancer prevalence in the UK up to the year 2040 were made under various scenarios of future incidence and survival. This chapter contains a detailed presentation and discussion of the results.

6.1 Background and literature

It is important to anticipate changes in cancer prevalence for the reasons set out in Chapter 4. But it is also of interest to explore the influence of each of the underlying elements that effect changes in cancer prevalence. Various factors should be considered: trends in incidence rates and survival, population growth and changes in the age structure of a population can each have a significant impact on cancer prevalence. In the literature, there are a number of examples of attempts to quantify each of these influences.

In general, cancer prevalence is increasing in Western countries. Stenbeck et al. (1999) used 35 years of national cancer registry data to show that cancer prevalence in Sweden increased between 1961 and 1995. By comparing, on a logarithmic scale, the observed trends in age-standardised incidence and age-standardised and crude prevalence, they estimated the proportion of the increases that could be attributable to trends in incidence rates and survival, population growth and population ageing. It was found that the largest proportion of observed increases in overall cancer prevalence was due to population dynamics – 40% and 47% for males and females, respectively. The second biggest influence was increasing cancer survival, accounting for 30% of the increases in cancer prevalence. The authors suggested that the increases in cancer prevalence were therefore primarily a consequence of increased life expectancy in the general population and better chances of surviving cancer – what they call "good forces".

Other studies have looked at individual types of cancer separately. Merrill (2001) found that the observed increases in prostate cancer prevalence in the USA between 1988 and 1997 were almost entirely due to increased incidence rates resulting from the introduction of PSA testing. Colonna et al. (2000) considered the ratio between prevalence proportions in France in 1992 and those in 1987, with and without age-standardisation, and concluded that ageing of the population was *not* one of the main factors acting to increase cancer prevalence. In another paper, some of the same authors claimed that changes in incidence rates and survival were the main factors

effecting changes in cancer prevalence in France between 1993 and 2002 (Colonna et al., 2008). In this later study, the influence of population dynamics was found to vary greatly across different cancer types; population ageing was responsible for increases in cancer prevalence of between 0.1% and 17.2%, and population growth was responsible for increases of between 4.7% and 13.5%.

De Angelis et al. (2009) projected breast cancer prevalence in the USA up to 2015 and gave specific consideration to the causes of geographical variation in prevalence across different states. The number of breast cancer survivors was projected to increase by 1 million between 2005 and 2015, from 2.4 million to 3.4 million. Breast cancer prevalence varied greatly between states but this was largely explained by the differing age structures of the population in each state as well as different state-specific incidence rates.

When producing projections of cancer prevalence, many authors have advocated using multiple different assumptions regarding future incidence rates and survival (Capocaccia et al., 1995, 1997; Verdecchia et al., 2002; Heinavaara and Hakulinen, 2006; Tabata et al., 2008). Primarily, the purpose of such a set of different assumptions is to provide a plausible range of estimates of future cancer prevalence, but they also provide a way of independently assessing the likely influences of underlying trends in incidence, survival or population demographics on future cancer prevalence.

The factors driving trends and regional variations in cancer prevalence largely depend on the specific circumstances under study (e.g. country, cancer type etc.). In this chapter, projections of cancer prevalence in the UK are presented. By considering different scenarios of future incidence rates and survival in detail, the influence of each is assessed independently from the influence of population dynamics.

6.2 Projection scenarios

Projections of cancer prevalence are highly dependent on the assumptions surrounding future incidence and survival that are made. In the previous chapter, regression procedures were described to provide estimates of future incidence and mortality under the assumption that existing trends in each will continue. These data provided input to the model that allowed the basic projections shown in this chapter to be made.

The model is, however, flexible in the sense that a wide range of assumptions can be used to specify the input data – the output is then the anticipated prevalence based on those assumptions. Long-term projections of cancer prevalence in the UK were sought

for the next three decades, up to the year 2040, a time period over which the assumption that existing trends will continue may not be realistic. For example, recent decades have seen advances in medicine that have, together with other factors, led to generally increasing survival for most types of cancer – it is perhaps optimistic to expect such increases to continue in the same way for the next 30 years. Other factors, such as the introduction of a breast cancer screening programme in the UK or the PSA test for prostate cancer, have caused rapid increases in recorded cancer incidence rates and changes to case-mix which cannot realistically be extrapolated to 2040 since they are clearly the result of specific interventions. By the same token, it is not possible to anticipate the effect of any new public health initiative or screening programme that might be introduced in the future.

For these reasons, it is desirable to provide a range of estimates of future cancer prevalence based on different assumptions regarding future incidence rates, survival and population demographics. Two different assumptions were used: a) the *dynamic* assumption (denoted ↕ in figures) that specified existing trends would continue in the period 2009–2040; and b) the *static* assumption (denoted ↔ in figures) that specified input data would remain constant from the most recent data year (2008) all the way to 2040. The model for projecting cancer prevalence was then run multiple times by applying each assumption to each of the inputs in turn, thereby defining a set of scenarios for which future cancer prevalence was estimated – see Table 6.1.

Table 6.1. Scenarios for which cancer prevalence was projected.

	Scenario number			
Input data	1*	2	3*	4
Incidence rates	Dynamic ↕	Static ↔	Dynamic ↕	Static ↔
Survival	Dynamic ↕	Static ↔	Static ↔	Dynamic ↕
Population demographics	Dynamic ↕	Dynamic ↕	Dynamic ↕	Dynamic ↕

*Incidence rates for prostate cancer were assumed to be static under scenarios 1 and 3, due to the unreliability of the projected prostate cancer incidence rates (see discussion in section 5.5.1). Scenarios 3 and 4 are therefore the same as scenarios 2 and 1, respectively, for prostate cancer, and as such are omitted from results tables and figures.

Population demographics were always assumed to be dynamic, based on projections provided by ONS (see section 5.2). These do, however, incorporate assumptions about general population mortality rates (including those for cancer survivors) which are separate from, and potentially inconsistent with, the static/dynamic assumptions used here.

The model for projecting cancer prevalence is specified in terms of mortality probabilities for survivors at different time points after diagnosis, rather than explicitly in terms of survival probabilities (see Chapter 4). Nonetheless, changes in one correspond directly to changes in the other and so, in practice, future survival scenarios were specified in terms of the mortality probabilities $\rho_{T,a}$ and $p_{T,y,a}$. These are defined in section 4.4.1.2.

Incidence rates and survival under the dynamic assumption were specified using the results described in section 5.5. Under the static assumption, age and sex-specific incidence rates and survival were calculated for the year 2008 and applied to all subsequent years.

6.3 Materials and methods

Cancer registry data for England were available covering diagnoses in the period 1971–2008, as described in section 5.2. The model was therefore able to produce estimates of limited duration 38-year cancer prevalence in England up to 2040, according to attained age, time since diagnosis, cancer type and sex, for each of the four scenarios in Table 6.1. These were then adjusted to account for survivors diagnosed more than 38 years previously (i.e. to give estimates of complete prevalence), using completeness indices previously calculated as part of the work described in Chapter 2 – see Table 6.2. It was assumed that these completeness indices (calculated based on prevalence data for 2005) applied for all years in the period 2009–2040; this was considered to be reasonable given that the required adjustments were small, although (as noted in section 2.4) completeness indices for male lung cancer may have been underestimated. In addition to counts, prevalence proportions (per 100,000 population) in broad age groups were also calculated using the same projected national population figures for England supplied by ONS.

Estimates for England were then generalised to the UK population by assuming that prevalence proportions for the UK were the same as those for England in each broad age group and time since diagnosis band – a reasonable assumption given that, currently, the population of England accounts for approximately 84% of the UK population (Office for National Statistics, 2011a).

Table 6.2. 38-year completeness indices for cancer prevalence in England, 2005, by broad age group, sex and cancer type.

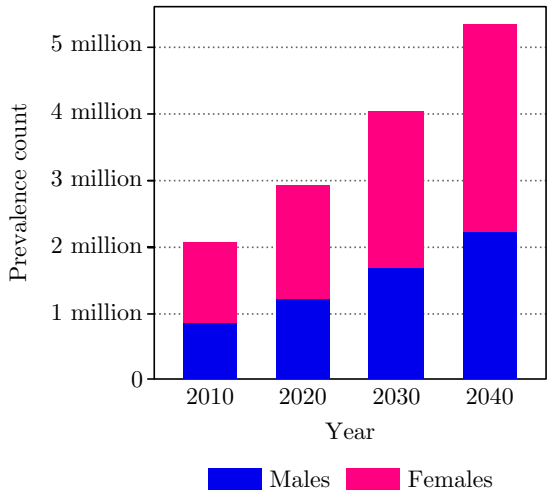
	Age group		
	0–44	45–64	≥65
Males			
Colon, rectum and anus	1.0000	0.9980	0.9795
Lung, bronchus and trachea*	1.0000	0.9955	0.8032
Prostate	1.0000	1.0000	0.9978
All other malignant neoplasms	0.9958	0.9672	0.9422
Females			
Colon, rectum and anus	0.9997	0.9946	0.9576
Lung, bronchus and trachea	1.0000	0.9976	0.9371
Breast	1.0000	0.9983	0.9214
All other malignant neoplasms	0.9970	0.9673	0.8757

*Possibly underestimated in older age groups, see discussion in section 2.4.

6.4 Results

Results are displayed in Figures 6.1–6.10 and tabulated in Tables 6.3–6.28. Figure 6.1 shows the projected increases in the total number of survivors of all malignant neoplasms (excluding non-melanoma skin cancer) in the UK under projection scenario 1. Figures 6.2–6.7 show the projected number of survivors (complete prevalence) in the UK from 2009 to 2040 under each of the scenarios described in Table 6.1, according to cancer type, sex and broad attained age group (0–44, 45–64 and ≥65 years). Due to the unrealistically high projections of prostate cancer incidence rates that were produced by the regression model (as discussed in section 5.5.1), the dynamic assumption was not used for prostate cancer incidence and therefore only two scenarios are included in Figure 6.4 and Tables 6.9 and 6.10.

Figure 6.1. Total number of survivors of all malignant neoplasms combined (excluding non-melanoma skin cancer) in the UK, 2010–2040, under projection scenario 1*.



*Dynamic incidence rates (for all cancer types except prostate), dynamic survival and dynamic population demographics.

Figures 6.8–6.10 display results under projection scenario 1. Figure 6.8 shows projected changes in the age distribution of cancer survivors for each sex and cancer type. Figure 6.9 shows projected changes in the proportion of survivors in each of the time since diagnosis groups <1, 1–5 and ≥5 years. Finally, Figure 6.10 shows projected changes in the proportion of the total number of survivors accounted for by each type of cancer.

Tables 6.3–6.28 are arranged in pairs – the first of each pair contains projected cancer prevalence counts and proportions (per 100,000 population) for the years 2010, 2020, 2030 and 2040; the second contains the corresponding average annual percentage change in cancer prevalence in the decades 2010–2020, 2020–2030 and 2030–2040. Tables 6.3 and 6.4 contain results for complete prevalence in all age groups combined under projection scenario 1. Tables 6.5–6.16 contain results for all projection scenarios, disaggregated according to attained age group. Finally, Tables 6.17–6.28 contain results for all projection scenarios, disaggregated according to time since diagnosis.

Considering the results for projection scenario 1, the total number of survivors in the UK (all malignant neoplasms combined) is estimated to grow by approximately 1 million every decade – from 2.1 million in 2010 to 2.9, 4.0 and 5.3 million in 2020, 2030 and 2040, respectively. Not only will there be more cancer survivors under this scenario, but they will account for a larger proportion of the population – from 2.8% of the male population in 2010 to 6.2% in 2040, and from 3.9% to 8.5% of the female population (Table 6.3). However, the growth rate of prevalence is projected to slow down over the

30 year period. The average annual percentage change in the number of survivors (proportion of the population) is projected to decrease from 3.7 (3.0) in the 2010s to 2.7 (2.4) in the 2040s for males, and from 3.3 (2.7) to 2.9 (2.6) for females – see Table 6.4.

The number of survivors, and the proportion of the population that they comprise, are also projected to increase substantially under this scenario for each of the individual cancer types studied, with one exception: male lung cancer prevalence is projected to exhibit only modest increases in terms of numbers (0.3%, 0.3% and 0.1% per year in each decade) and to decrease slightly in terms of the proportion of the population (-0.5%, -0.3% and -0.2% per year in each decade) – see Table 6.4. This is in contrast to the prevalence of female lung cancer which is projected to increase by 4.7%, 4.6% and 4.1% (count) and by 4.0%, 4.0% and 3.8% (proportion of the population) in each decade – these are the largest projected proportional increases of any cancer apart from prostate cancer in the 2010s. Despite this, however, the number of female lung cancer survivors will remain relatively modest under this scenario, accounting for just 3.1% of all female cancer survivors by 2040 (Figure 6.10). It is prostate cancer prevalence which is projected to increase at the fastest rate among males, despite this scenario assuming static prostate cancer incidence rates from 2009 onwards – by 2040 the total number of prostate cancer survivors will have more than trebled to around 830,000, accounting for approximately 2.3% of the male population overall (and considerably more than this in the older age groups).

Currently, a large majority of cancer survivors are aged 65 years and over. In 2009, the proportion was 66.7% among male cancer survivors and 59.4% among females. Under projection scenario 1 this will rise to 82.3% and 73.1% in 2040, for males and females respectively (Figure 6.8). In the youngest age group, 0–44 years, the proportion of the population who are cancer survivors (all malignant neoplasms combined) is projected to increase modestly between 2010 and 2040 – from just under 0.4% to just over 0.4% for males, and from 0.5% to 0.7% for females. In the age group 45–64 years these increases are also relatively modest, from 2.7% to 3.6% of the male population and from 5.0% to 8.4% of the female population. However, more dramatic increases are projected for the oldest age group; the proportion of the population aged at least 65 years who are cancer survivors will almost double to 23.3% (males) and 24.9% (females) by 2040 under this scenario (Table 6.15). By far the largest contributing cancer types in this age group are prostate and female breast – under this projection scenario, in 2040, 10.0% of

the male population will be prostate cancer survivors (Table 6.9), and 13.3% of the female population will be breast cancer survivors (Table 6.11).

Still considering the same projection scenario, Figure 6.9 shows the projected changes in the distribution of cancer survivors between different time since diagnosis bands. A large majority of all cancer survivors are currently at least five years beyond diagnosis, and this proportion is projected to increase from 55.2% in 2009 to 65.5% in 2040 for males and from 66.0% to 70.9% for females. Accordingly, the proportion of all survivors who are less than five years beyond diagnosis is projected to decrease. Nonetheless, the actual number of survivors who are less than five years beyond diagnosis will more than double from 2010 to 2040, from around 800,000 to around 1.7 million (Table 6.27). Projected prevalence counts for individual cancer types show similar patterns. Lung, bronchus and trachea cancer is the only type studied here for which the number of survivors who were less than one year beyond diagnosis in 2009 exceeded the number who were between one and five years beyond diagnosis. For males, this is projected to remain true for the whole period up to 2040 (Table 6.19), but for females the number of lung cancer survivors between one and five years beyond diagnosis is projected to surpass the number less than one year beyond diagnosis from 2026 onwards. The number of prostate cancer survivors who are at least five years beyond diagnosis will reach 578,000 in 2040 under this scenario and will account for 69.6% of all prostate cancer survivors, a sizeable increase compared with the proportion in 2009 (42.2%).

Scenario 1 (i.e. dynamic incidence rates, survival and population demographics) generally resulted in the highest projected cancer prevalence, and scenario 2 (i.e. static incidence rates and survival, dynamic population demographics) in the lowest. For all malignant neoplasms and ages combined, the number of cancer survivors in 2040 was projected to be 5.3 million under scenario 1, and 3.5 million under scenario 2 (Table 6.15). By comparison, scenarios 3 and 4 gave projections of 4.1 million and 4.5 million, respectively.

All projection scenarios for the majority of cancer types, sexes and age groups resulted in generally increasing cancer prevalence. The notable exception was, however, male lung cancer prevalence which exhibited quite different patterns to that of the other cancer types studied (Figure 6.3 and Table 6.7). In the youngest age group (0–44 years) it was scenario 2 that resulted in the highest projected male lung cancer prevalence, although the number of survivors in this age group is very small (less than 1,000). In the age groups 45–64 and ≥65 years, scenario 4 (i.e. dynamic survival and population demographics) gave the highest projected prevalence and scenario 3 (i.e. dynamic

incidence rates and population demographics) the lowest. Overall, for all age groups combined, the number of male lung cancer survivors is projected to remain roughly constant under projection scenario 1 (i.e. dynamic incidence rates, survival and population demographics).

In most cases, each projection scenario resulted in substantially different prevalence projections. Perhaps the most notable exception was for female colon, rectum and anus cancer. In the age groups 45–64 and ≥65 years, there was little difference between the projections made under scenarios 2 and 3, and similarly between those made under scenarios 1 and 4.

Figure 6.2. Complete prevalence of colon, rectum and anus cancer in the UK, 2009–2040, by attained age, sex and projection scenario.

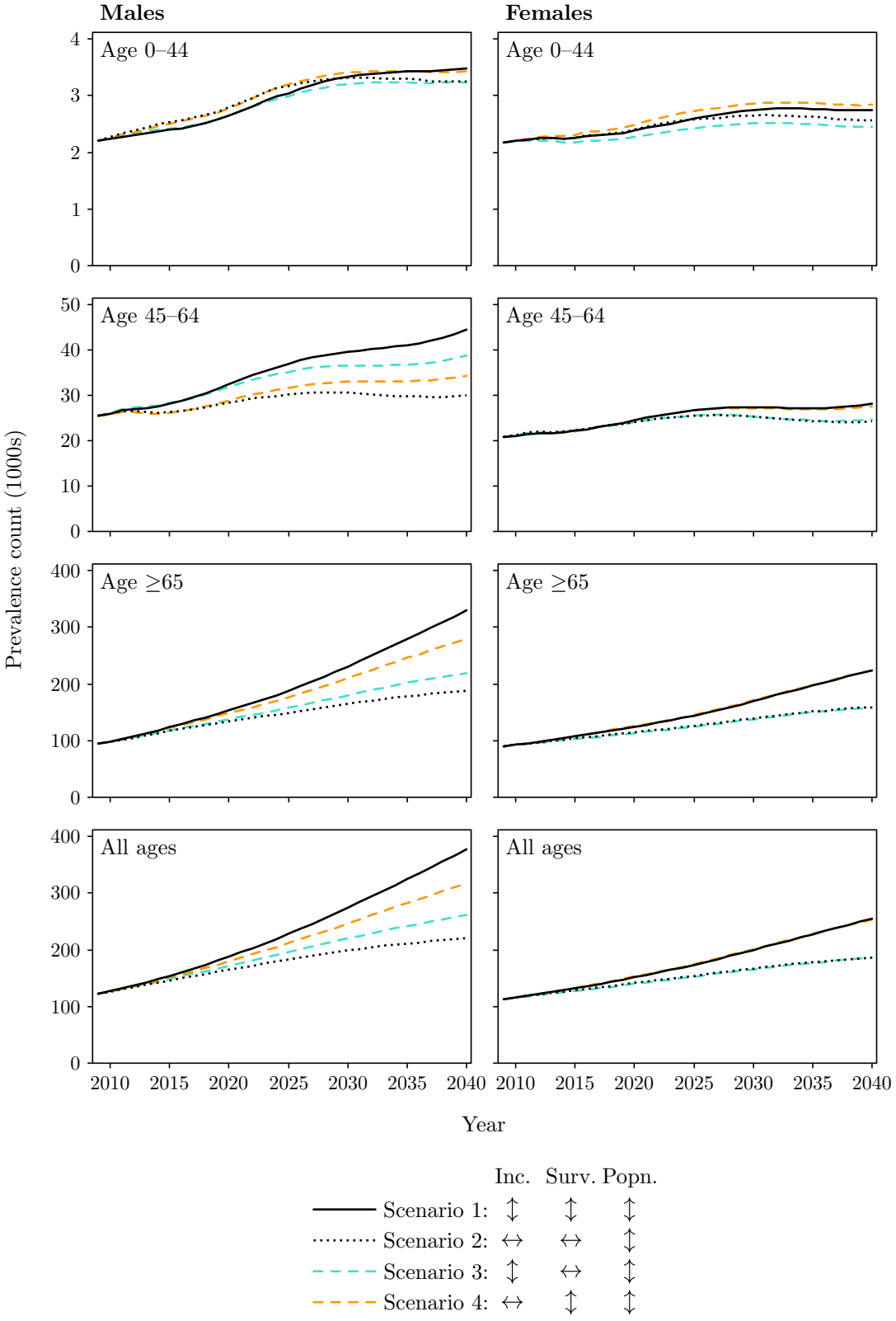


Figure 6.3. Complete prevalence of lung, bronchus and trachea cancer in the UK, 2009–2040, by attained age, sex and projection scenario.

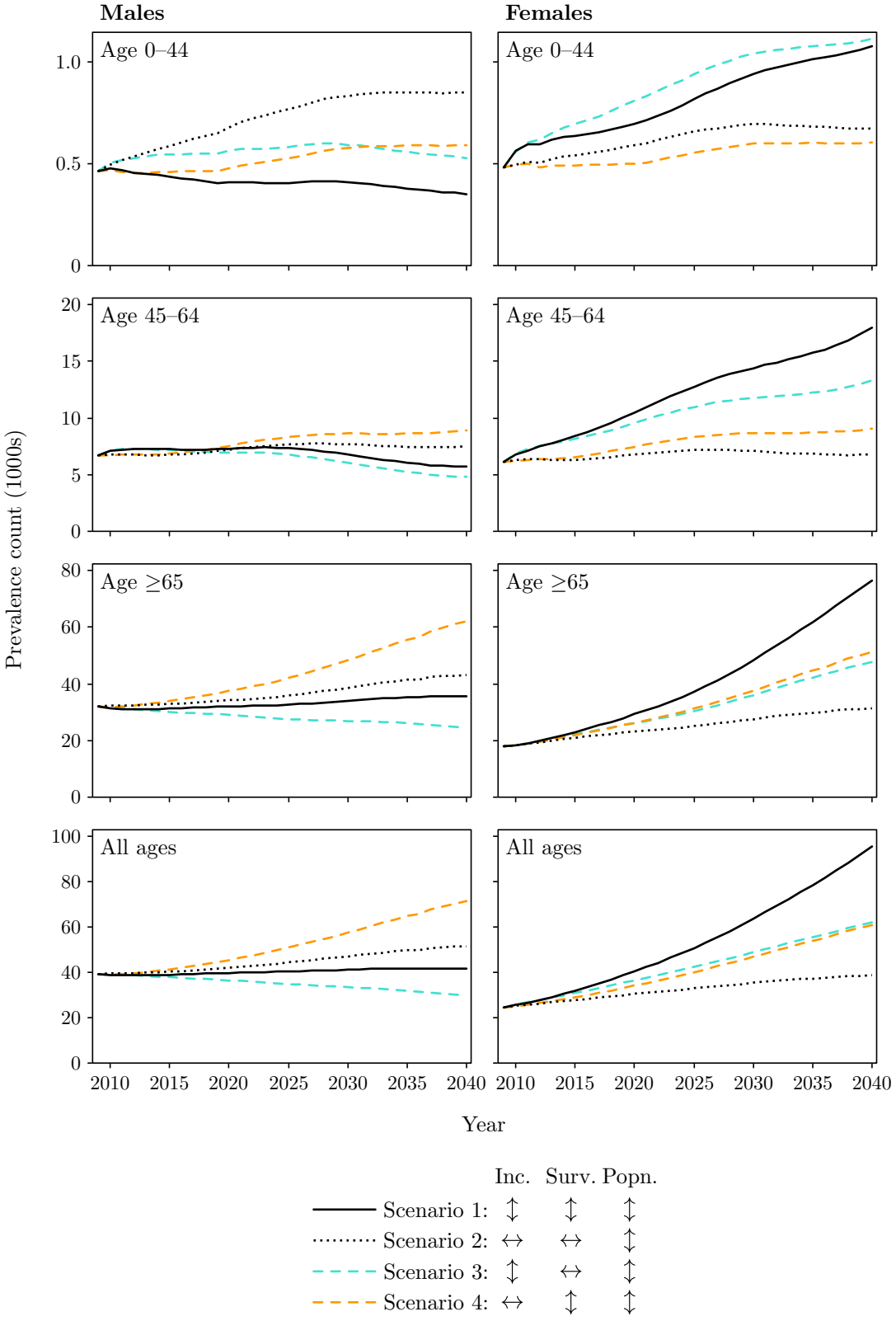
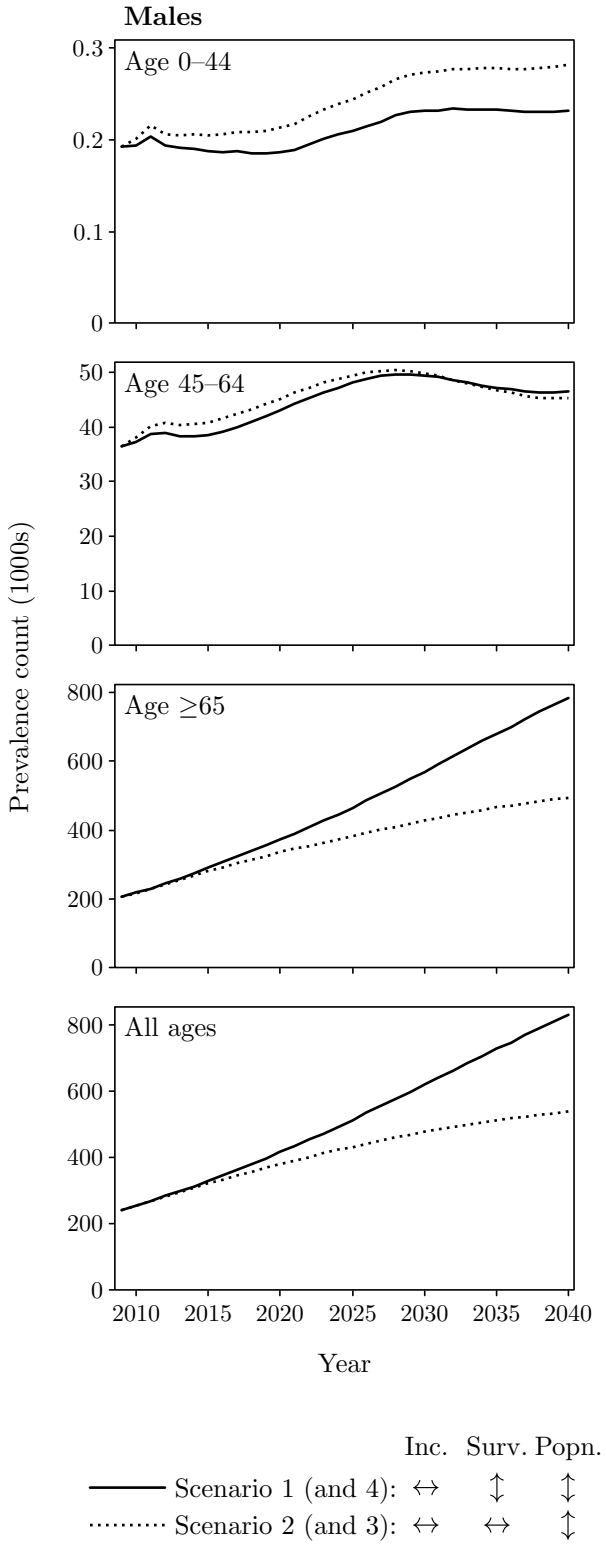


Figure 6.4. Complete prevalence of prostate cancer in the UK, 2009–2040, by attained age and projection scenario*.



*Scenarios 3 and 4 are omitted for prostate cancer due to being the same as scenarios 2 and 1, respectively; see Table 6.1 for details.

Figure 6.5. Complete prevalence of female breast cancer in the UK, 2009–2040, by attained age and projection scenario.

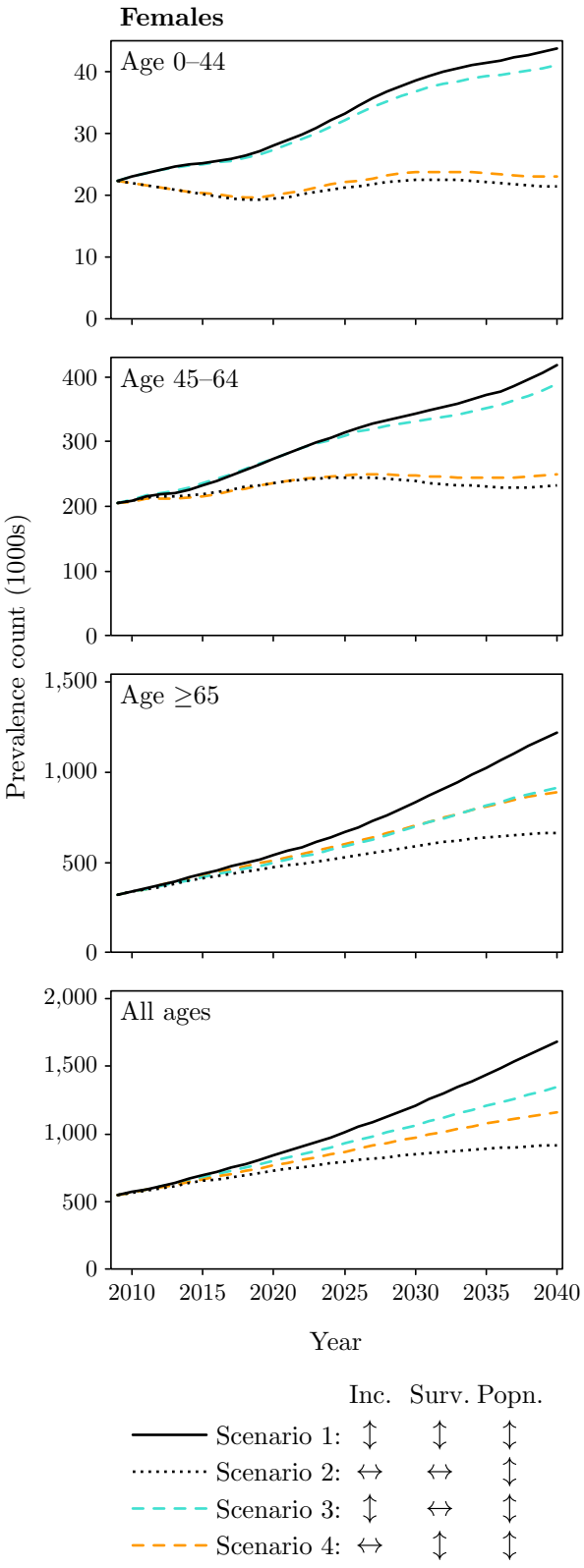


Figure 6.6. Complete prevalence of all other malignant neoplasms (excluding non-melanoma skin cancer) in the UK, 2009–2040, by attained age, sex and projection scenario.

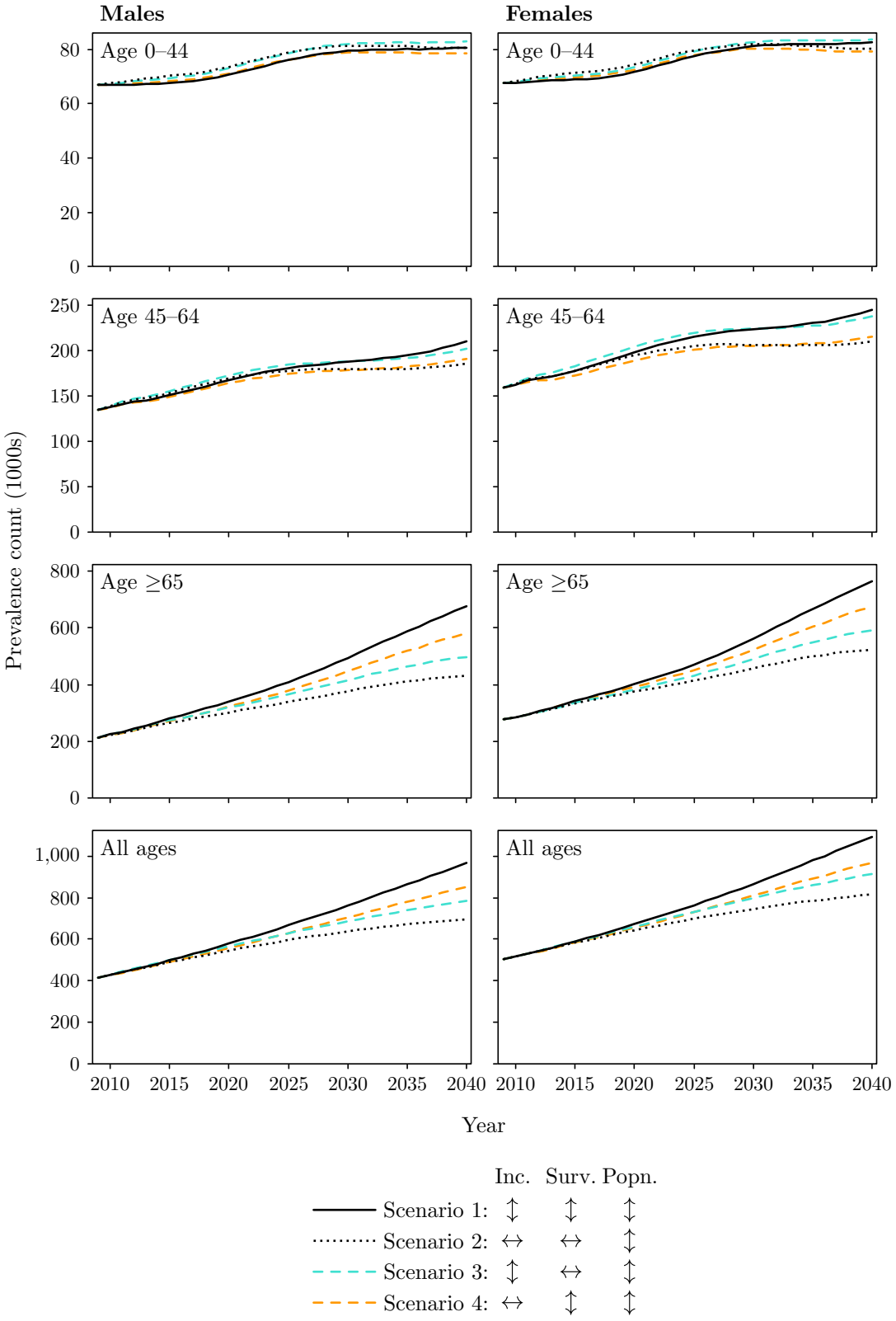


Figure 6.7. Complete prevalence of all malignant neoplasms (excluding non-melanoma skin cancer) in the UK, 2009–2040, by attained age, sex and projection scenario.

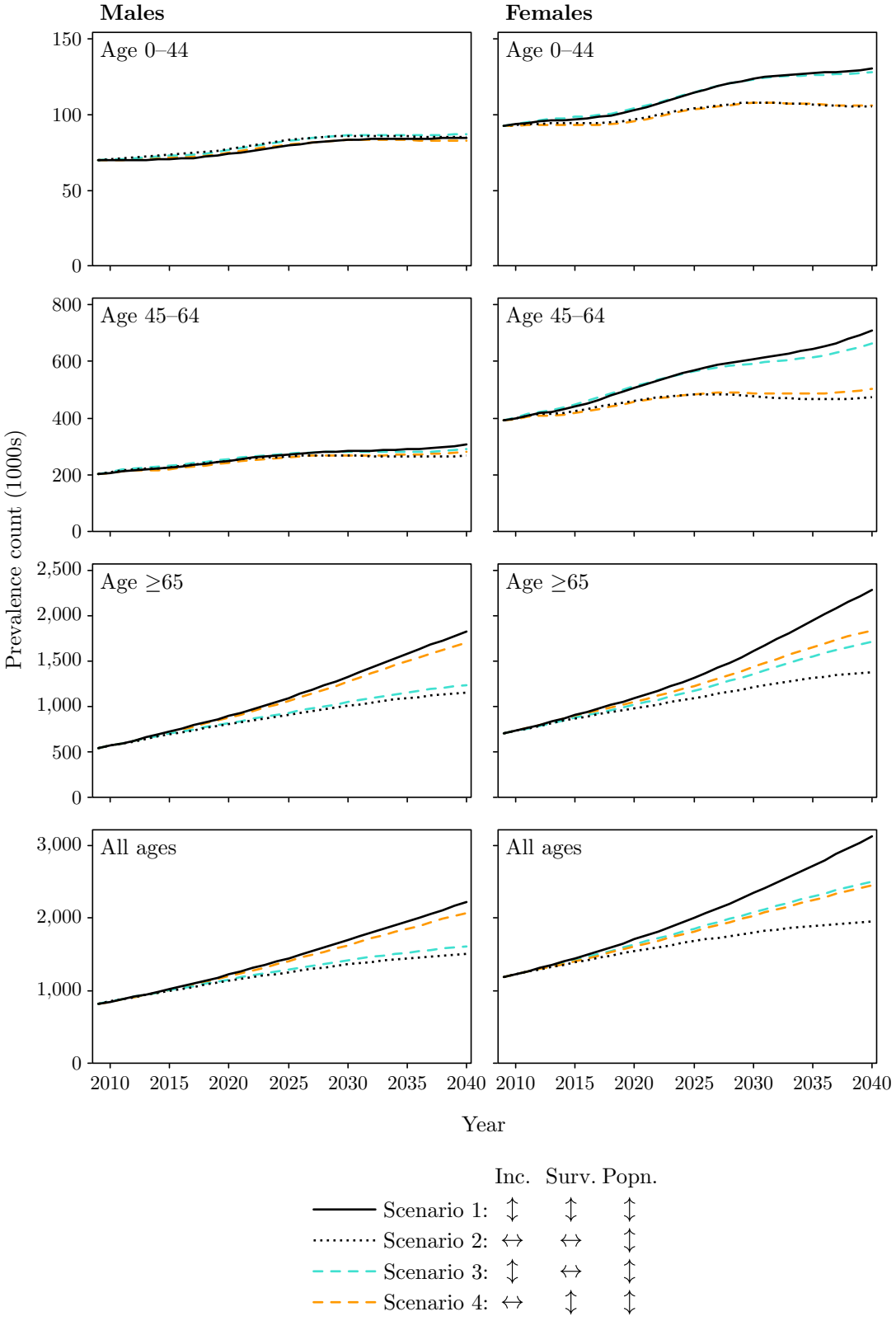
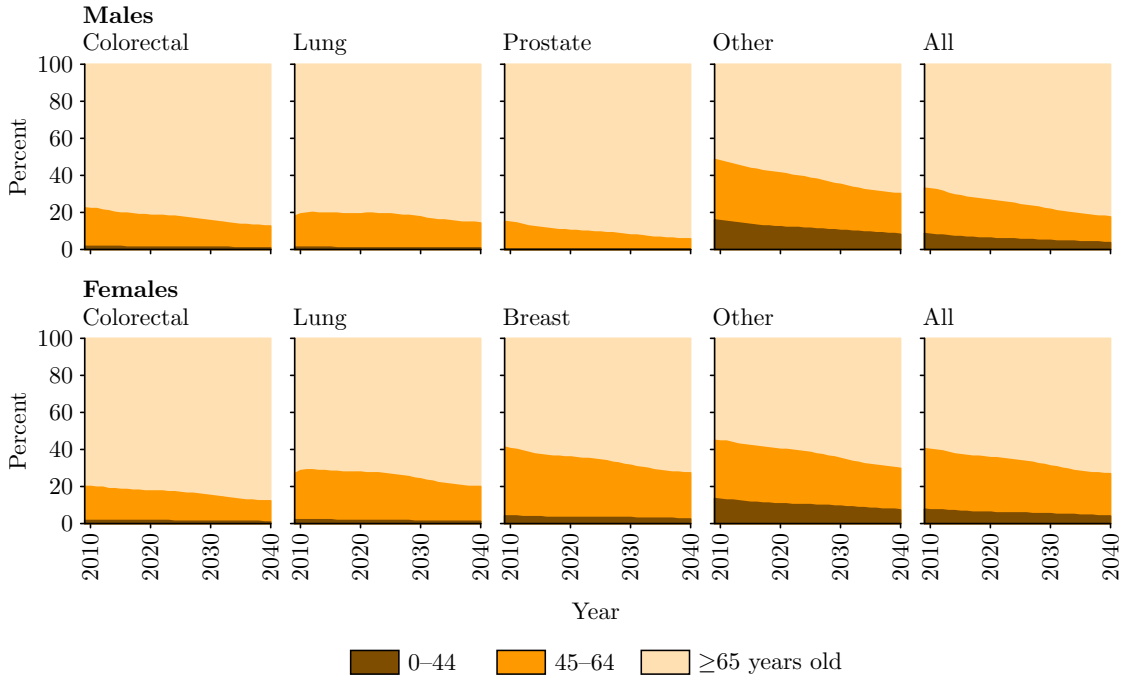
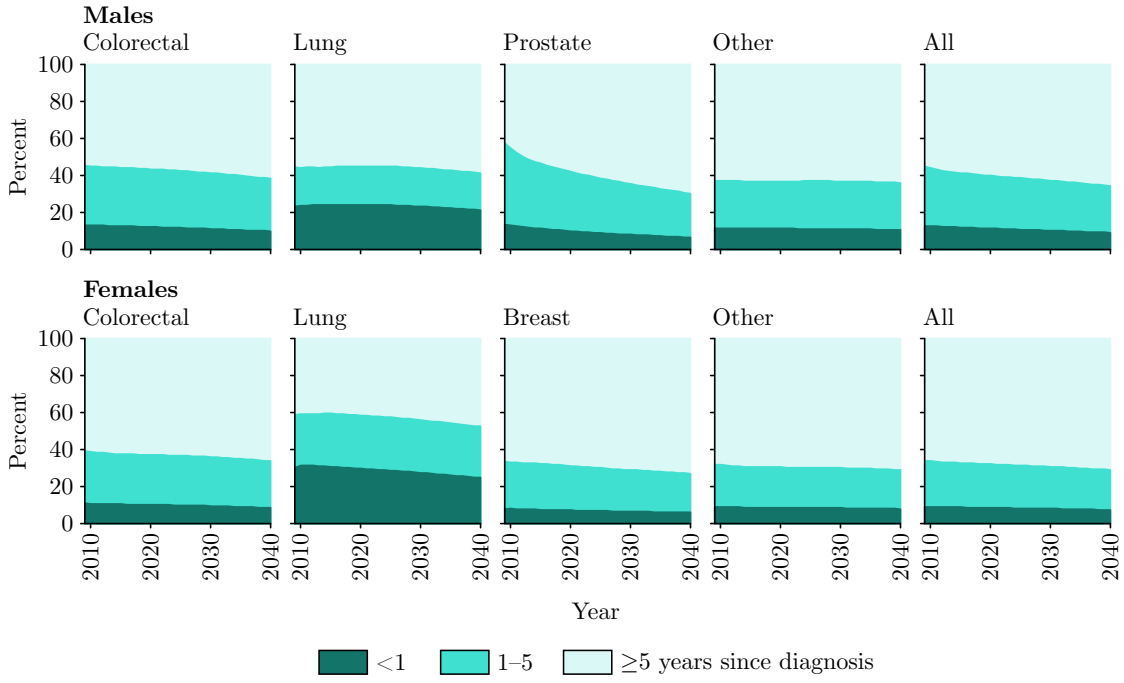


Figure 6.8. Age distribution of cancer survivors in the UK, 2009–2040, by cancer type* and sex, under projection scenario 1†. Proportion of total number of survivors in each attained age group.



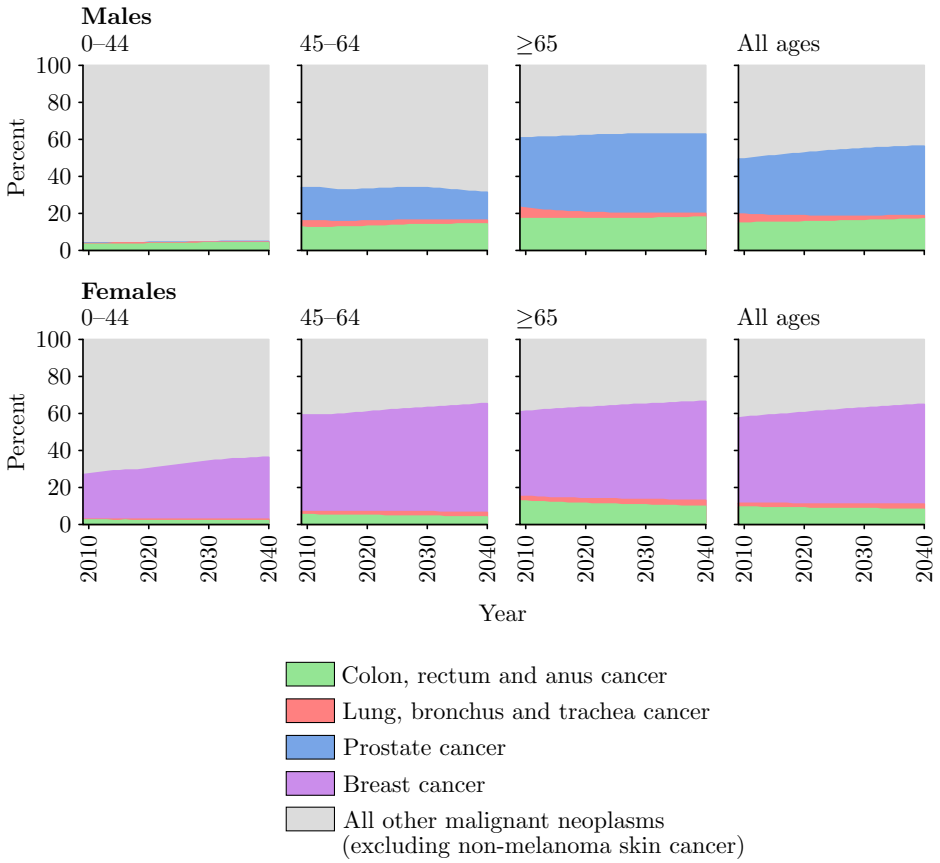
*ICD-10 codes as follows: Colorectal = C18–C21; Lung = C33–C34; Prostate = C61; Breast = C50; Other = C00–C97 excluding C44 and those mentioned previously; All = C00–C96 excluding C44. †Dynamic incidence rates (for all cancer types except prostate), dynamic survival and dynamic population demographics.

Figure 6.9. Time since diagnosis distribution of cancer survivors in the UK, 2009–2040, by cancer type* and sex, under projection scenario 1†. Proportion of total number of survivors in each time since diagnosis band.



*ICD-10 codes as follows: Colorectal = C18–C21; Lung = C33–C34; Prostate = C61; Breast = C50; Other = C00–C97 excluding C44 and those mentioned previously; All = C00–C97 excluding C44. †Dynamic incidence rates (for all cancer types except prostate), dynamic survival and dynamic population demographics.

Figure 6.10. Distribution between cancer types of cancer survivors in the UK, 2009–2040, by attained age and sex, under projection scenario 1*. Proportion of total number of survivors accounted for by each cancer type.



*Dynamic incidence rates (for all cancer types except prostate), dynamic survival and dynamic population demographics.

Table 6.3. Complete prevalence in the UK, 2010–2040, by cancer type* and sex, under projection scenario 1†. Number of survivors (proportion of the population per 100,000).

	Year			
	2010	2020	2030	2040
Males				
Colorectal	127,076 (415)	188,354 (572)	273,966 (783)	377,200 (1,048)
Lung	38,827 (127)	39,825 (121)	41,208 (118)	41,612 (116)
Prostate	255,431 (835)	416,104 (1,264)	619,544 (1,771)	830,533 (2,306)
Other	428,680 (1,401)	578,810 (1,759)	761,941 (2,178)	966,396 (2,684)
All	850,014 (2,777)	1,223,093 (3,717)	1,696,659 (4,850)	2,215,742 (6,153)
Females				
Colorectal	116,439 (368)	151,599 (451)	199,710 (561)	254,956 (697)
Lung	25,649 (81)	40,477 (120)	63,607 (179)	95,381 (261)
Breast	569,883 (1,803)	840,460 (2,500)	1,212,319 (3,406)	1,682,737 (4,598)
Other	516,737 (1,635)	672,070 (1,999)	866,271 (2,434)	1,091,560 (2,983)
All	1,228,708 (3,887)	1,704,606 (5,071)	2,341,907 (6,579)	3,124,634 (8,538)

*ICD-10 codes as follows: Colorectal = C18–C21; Lung = C33–C34; Prostate = C61; Breast = C50; Other = C00–C97 excluding C44 and those mentioned previously; All = C00–C96 excluding C44. †Dynamic incidence rates (for all cancer types except prostate), dynamic survival and dynamic population demographics.

Table 6.4. Rate of change of complete prevalence in the UK, 2010–2040, by cancer type* and sex, under projection scenario 1†. Average annual percentage change in number of survivors (proportion of the population).

	Period		
	2010–2020	2020–2030	2030–2040
Males			
Colorectal	4.0 (3.3)	3.8 (3.2)	3.2 (3.0)
Lung	0.3 (-0.5)	0.3 (-0.3)	0.1 (-0.2)
Prostate	5.0 (4.2)	4.1 (3.4)	3.0 (2.7)
Other	3.0 (2.3)	2.8 (2.2)	2.4 (2.1)
All	3.7 (3.0)	3.3 (2.7)	2.7 (2.4)
Females			
Colorectal	2.7 (2.0)	2.8 (2.2)	2.5 (2.2)
Lung	4.7 (4.0)	4.6 (4.0)	4.1 (3.8)
Breast	4.0 (3.3)	3.7 (3.1)	3.3 (3.0)
Other	2.7 (2.0)	2.6 (2.0)	2.3 (2.1)
All	3.3 (2.7)	3.2 (2.6)	2.9 (2.6)

*ICD-10 codes as follows: Colorectal = C18–C21; Lung = C33–C34; Prostate = C61; Breast = C50; Other = C00–C97 excluding C44 and those mentioned previously; All = C00–C96 excluding C44. †Dynamic incidence rates (for all cancer types except prostate), dynamic survival and dynamic population demographics.

Table 6.5. Complete prevalence of colon, rectum and anus cancer in the UK, 2010–2040, by attained age, sex and projection scenario. Number of survivors (proportion of the population per 100,000).

		Year							
Age	Scenario	2010		2020		2030		2040	
Males									
0–44	1	2,236	(12)	2,652	(14)	3,334	(17)	3,482	(18)
	2	2,270	(12)	2,798	(15)	3,312	(17)	3,243	(16)
	3	2,249	(12)	2,655	(14)	3,200	(16)	3,238	(16)
	4	2,257	(12)	2,784	(15)	3,406	(17)	3,429	(17)
45–64	1	25,925	(333)	32,402	(393)	39,608	(483)	44,511	(527)
	2	25,938	(333)	28,364	(344)	30,541	(373)	30,018	(355)
	3	26,106	(335)	31,780	(386)	36,528	(446)	38,739	(459)
	4	25,759	(331)	28,876	(351)	33,020	(403)	34,354	(407)
≥65	1	98,915	(2,186)	153,301	(2,653)	231,024	(3,248)	329,207	(4,205)
	2	98,133	(2,169)	133,966	(2,318)	164,931	(2,319)	187,377	(2,394)
	3	98,338	(2,174)	137,774	(2,384)	180,276	(2,535)	219,793	(2,808)
	4	98,689	(2,181)	148,401	(2,568)	210,082	(2,954)	279,462	(3,570)
All	1	127,076	(415)	188,354	(572)	273,966	(783)	377,200	(1,048)
	2	126,342	(413)	165,128	(502)	198,784	(568)	220,637	(613)
	3	126,693	(414)	172,209	(523)	220,005	(629)	261,769	(727)
	4	126,704	(414)	180,061	(547)	246,508	(705)	317,245	(881)
Females									
0–44	1	2,202	(12)	2,384	(13)	2,751	(15)	2,745	(15)
	2	2,199	(12)	2,401	(13)	2,653	(14)	2,560	(14)
	3	2,195	(12)	2,274	(13)	2,511	(13)	2,448	(13)
	4	2,205	(12)	2,488	(14)	2,865	(15)	2,835	(15)
45–64	1	21,067	(261)	24,493	(286)	27,397	(329)	28,237	(335)
	2	21,317	(264)	24,199	(283)	25,273	(303)	24,244	(288)
	3	21,180	(263)	24,133	(282)	25,409	(305)	24,752	(294)
	4	21,203	(263)	24,552	(287)	27,217	(327)	27,577	(327)
≥65	1	93,171	(1,621)	124,722	(1,811)	169,562	(2,028)	223,974	(2,436)
	2	92,716	(1,613)	114,998	(1,670)	139,449	(1,668)	158,976	(1,729)
	3	92,638	(1,611)	113,594	(1,650)	137,860	(1,649)	158,604	(1,725)
	4	93,248	(1,622)	125,944	(1,829)	170,928	(2,045)	223,398	(2,430)
All	1	116,439	(368)	151,599	(451)	199,710	(561)	254,956	(697)
	2	116,232	(368)	141,599	(421)	167,374	(470)	185,780	(508)
	3	116,013	(367)	140,000	(416)	165,780	(466)	185,804	(508)
	4	116,656	(369)	152,984	(455)	201,010	(565)	253,810	(694)

	Inc.	Surv.	Popn.
Scenario 1:	↕	↕	↕
Scenario 2:	↔	↔	↕
Scenario 3:	↕	↔	↕
Scenario 4:	↔	↕	↕

Table 6.6. Rate of change of complete prevalence of colon, rectum and anus cancer in the UK, 2010–2040, by attained age, sex and projection scenario. Average annual percentage change in number of survivors (proportion of the population).

		Period		
Age	Scenario	2010–2020	2020–2030	2030–2040
Males				
0–44	1	1.7 (1.4)	2.3 (1.9)	0.4 (0.4)
	2	2.1 (1.8)	1.7 (1.3)	-0.2 (-0.2)
	3	1.7 (1.4)	1.9 (1.5)	0.1 (0.1)
	4	2.1 (1.8)	2.0 (1.6)	0.1 (0.0)
45–64	1	2.3 (1.7)	2.0 (2.1)	1.2 (0.9)
	2	0.9 (0.3)	0.7 (0.8)	-0.2 (-0.5)
	3	2.0 (1.4)	1.4 (1.5)	0.6 (0.3)
	4	1.1 (0.6)	1.4 (1.4)	0.4 (0.1)
≥65	1	4.5 (2.0)	4.2 (2.0)	3.6 (2.6)
	2	3.2 (0.7)	2.1 (0.0)	1.3 (0.3)
	3	3.4 (0.9)	2.7 (0.6)	2.0 (1.0)
	4	4.2 (1.6)	3.5 (1.4)	2.9 (1.9)
All	1	4.0 (3.3)	3.8 (3.2)	3.2 (3.0)
	2	2.7 (2.0)	1.9 (1.3)	1.0 (0.8)
	3	3.1 (2.4)	2.5 (1.9)	1.8 (1.5)
	4	3.6 (2.8)	3.2 (2.6)	2.6 (2.3)
Females				
0–44	1	0.8 (0.6)	1.4 (1.1)	0.0 (0.0)
	2	0.9 (0.7)	1.0 (0.6)	-0.4 (-0.4)
	3	0.4 (0.2)	1.0 (0.6)	-0.3 (-0.3)
	4	1.2 (1.0)	1.4 (1.0)	-0.1 (-0.1)
45–64	1	1.5 (0.9)	1.1 (1.4)	0.3 (0.2)
	2	1.3 (0.7)	0.4 (0.7)	-0.4 (-0.5)
	3	1.3 (0.7)	0.5 (0.8)	-0.3 (-0.4)
	4	1.5 (0.9)	1.0 (1.3)	0.1 (0.0)
≥65	1	3.0 (1.1)	3.1 (1.1)	2.8 (1.8)
	2	2.2 (0.4)	1.9 (0.0)	1.3 (0.4)
	3	2.1 (0.2)	2.0 (0.0)	1.4 (0.5)
	4	3.1 (1.2)	3.1 (1.1)	2.7 (1.7)
All	1	2.7 (2.0)	2.8 (2.2)	2.5 (2.2)
	2	2.0 (1.4)	1.7 (1.1)	1.0 (0.8)
	3	1.9 (1.3)	1.7 (1.1)	1.1 (0.9)
	4	2.7 (2.1)	2.8 (2.2)	2.4 (2.1)

	Inc.	Surv.	Popn.
Scenario 1:	↕	↕	↕
Scenario 2:	↔	↔	↕
Scenario 3:	↕	↔	↕
Scenario 4:	↔	↕	↕

Table 6.7. Complete prevalence of lung, bronchus and trachea cancer in the UK, 2010–2040, by attained age, sex and projection scenario. Number of survivors (proportion of the population per 100,000).

		Year			
Age	Scenario	2010	2020	2030	2040
Males					
0–44	1	476 (3)	410 (2)	408 (2)	351 (2)
	2	497 (3)	677 (4)	832 (4)	847 (4)
	3	505 (3)	561 (3)	591 (3)	529 (3)
	4	467 (3)	479 (3)	575 (3)	590 (3)
45–64	1	7,076 (91)	7,315 (89)	6,790 (83)	5,693 (67)
	2	6,753 (87)	7,201 (87)	7,697 (94)	7,515 (89)
	3	7,128 (92)	6,989 (85)	6,071 (74)	4,808 (57)
	4	6,709 (86)	7,550 (92)	8,639 (105)	8,940 (106)
≥65	1	31,275 (691)	32,101 (556)	34,011 (478)	35,568 (454)
	2	32,299 (714)	34,231 (592)	38,653 (543)	43,120 (551)
	3	31,639 (699)	29,036 (502)	26,952 (379)	24,594 (314)
	4	31,882 (705)	37,542 (650)	48,354 (680)	62,087 (793)
All	1	38,827 (127)	39,825 (121)	41,208 (118)	41,612 (116)
	2	39,549 (129)	42,109 (128)	47,182 (135)	51,482 (143)
	3	39,271 (128)	36,586 (111)	33,614 (96)	29,931 (83)
	4	39,059 (128)	45,571 (138)	57,568 (165)	71,618 (199)
Females					
0–44	1	561 (3)	695 (4)	938 (5)	1,073 (6)
	2	497 (3)	590 (3)	694 (4)	671 (4)
	3	558 (3)	808 (4)	1,038 (6)	1,111 (6)
	4	494 (3)	500 (3)	598 (3)	602 (3)
45–64	1	6,756 (84)	10,498 (123)	14,410 (173)	17,939 (213)
	2	6,262 (78)	6,791 (79)	7,097 (85)	6,784 (81)
	3	6,852 (85)	9,552 (112)	11,753 (141)	13,309 (158)
	4	6,180 (77)	7,468 (87)	8,666 (104)	9,070 (108)
≥65	1	18,331 (319)	29,284 (425)	48,259 (577)	76,369 (831)
	2	18,486 (322)	23,138 (336)	27,601 (330)	31,320 (341)
	3	18,450 (321)	26,195 (380)	36,088 (432)	47,695 (519)
	4	18,386 (320)	26,261 (381)	37,638 (450)	51,315 (558)
All	1	25,649 (81)	40,477 (120)	63,607 (179)	95,381 (261)
	2	25,244 (80)	30,518 (91)	35,392 (99)	38,774 (106)
	3	25,859 (82)	36,555 (109)	48,879 (137)	62,115 (170)
	4	25,061 (79)	34,229 (102)	46,902 (132)	60,987 (167)

	Inc.	Surv.	Popn.
Scenario 1:	↕	↕	↕
Scenario 2:	↔	↔	↕
Scenario 3:	↕	↔	↕
Scenario 4:	↔	↕	↕

Table 6.8. Rate of change of complete prevalence of lung, bronchus and trachea cancer in the UK, 2010–2040, by attained age, sex and projection scenario. Average annual percentage change in number of survivors (proportion of the population).

		Period		
Age	Scenario	2010–2020	2020–2030	2030–2040
Males				
0–44	1	-1.5 (-1.8)	0.0 (-0.4)	-1.5 (-1.5)
	2	3.1 (2.8)	2.1 (1.7)	0.2 (0.2)
	3	1.1 (0.7)	0.5 (0.1)	-1.1 (-1.1)
	4	0.2 (-0.1)	1.9 (1.4)	0.3 (0.2)
45–64	1	0.3 (-0.2)	-0.7 (-0.7)	-1.7 (-2.0)
	2	0.6 (0.1)	0.7 (0.7)	-0.2 (-0.5)
	3	-0.2 (-0.8)	-1.4 (-1.3)	-2.3 (-2.6)
	4	1.2 (0.6)	1.4 (1.4)	0.3 (0.0)
≥65	1	0.3 (-2.2)	0.6 (-1.5)	0.4 (-0.5)
	2	0.6 (-1.8)	1.2 (-0.9)	1.1 (0.1)
	3	-0.9 (-3.3)	-0.7 (-2.8)	-0.9 (-1.9)
	4	1.6 (-0.8)	2.6 (0.5)	2.5 (1.6)
All	1	0.3 (-0.5)	0.3 (-0.3)	0.1 (-0.2)
	2	0.6 (-0.1)	1.1 (0.5)	0.9 (0.6)
	3	-0.7 (-1.4)	-0.8 (-1.4)	-1.2 (-1.4)
	4	1.6 (0.8)	2.4 (1.7)	2.2 (1.9)
Females				
0–44	1	2.2 (2.0)	3.0 (2.6)	1.4 (1.3)
	2	1.7 (1.5)	1.6 (1.2)	-0.3 (-0.4)
	3	3.8 (3.6)	2.5 (2.1)	0.7 (0.7)
	4	0.1 (-0.1)	1.8 (1.4)	0.1 (0.1)
45–64	1	4.5 (3.9)	3.2 (3.5)	2.2 (2.1)
	2	0.8 (0.2)	0.4 (0.7)	-0.5 (-0.6)
	3	3.4 (2.8)	2.1 (2.4)	1.3 (1.1)
	4	1.9 (1.3)	1.5 (1.8)	0.5 (0.3)
≥65	1	4.8 (2.9)	5.1 (3.1)	4.7 (3.7)
	2	2.3 (0.4)	1.8 (-0.2)	1.3 (0.3)
	3	3.6 (1.7)	3.3 (1.3)	2.8 (1.9)
	4	3.6 (1.8)	3.7 (1.7)	3.1 (2.2)
All	1	4.7 (4.0)	4.6 (4.0)	4.1 (3.8)
	2	1.9 (1.3)	1.5 (0.9)	0.9 (0.6)
	3	3.5 (2.9)	2.9 (2.4)	2.4 (2.1)
	4	3.2 (2.5)	3.2 (2.6)	2.7 (2.4)

	Inc.	Surv.	Popn.
Scenario 1:	↕	↕	↕
Scenario 2:	↔	↔	↕
Scenario 3:	↕	↔	↕
Scenario 4:	↔	↕	↕

Table 6.9. Complete prevalence of prostate cancer in the UK, 2010–2040, by attained age and projection scenario*. Number of survivors (proportion of the population per 100,000).

		Year							
Age	Scenario	2010		2020		2030		2040	
0–44	1	193	(1)	187	(1)	232	(1)	232	(1)
	2	200	(1)	213	(1)	273	(1)	281	(1)
	3								
	4								
45–64	1	37,291	(479)	43,016	(522)	49,459	(604)	46,470	(550)
	2	38,103	(490)	45,157	(548)	49,858	(609)	45,356	(537)
	3								
	4								
≥65	1	217,947	(4,817)	372,901	(6,453)	569,854	(8,012)	783,831	(10,013)
	2	217,736	(4,813)	335,265	(5,802)	427,151	(6,006)	492,931	(6,297)
	3								
	4								
All	1	255,431	(835)	416,104	(1,264)	619,544	(1,771)	830,533	(2,306)
	2	256,039	(836)	380,635	(1,157)	477,282	(1,364)	538,568	(1,496)
	3								
	4								

*Scenarios 3 and 4 are omitted for prostate cancer due to being the same as scenarios 2 and 1, respectively; see Table 6.1 for details.

	Inc.	Surv.	Popn.
Scenario 1:	↔	↕	↕
Scenario 2:	↔	↔	↕

Table 6.10. Rate of change of complete prevalence of prostate cancer in the UK, 2010–2040, by attained age and projection scenario*. Average annual percentage change in number of survivors (proportion of the population).

		Period		
Age	Scenario	2010–2020	2020–2030	2030–2040
0–44	1	-0.4 (-0.7)	2.2 (1.8)	0.0 (0.0)
	2	0.6 (0.3)	2.5 (2.1)	0.3 (0.3)
	3			
	4			
45–64	1	1.4 (0.9)	1.4 (1.5)	-0.6 (-0.9)
	2	1.7 (1.1)	1.0 (1.1)	-0.9 (-1.2)
	3			
	4			
≥65	1	5.5 (3.0)	4.3 (2.2)	3.2 (2.3)
	2	4.4 (1.9)	2.5 (0.3)	1.4 (0.5)
	3			
	4			
All	1	5.0 (4.2)	4.1 (3.4)	3.0 (2.7)
	2	4.0 (3.3)	2.3 (1.7)	1.2 (0.9)
	3			
	4			

*Scenarios 3 and 4 are omitted for prostate cancer due to being the same as scenarios 2 and 1, respectively; see Table 6.1 for details.

	Inc.	Surv.	Popn.
Scenario 1:	↔	↕	↕
Scenario 2:	↔	↔	↕

Table 6.11. Complete prevalence of female breast cancer in the UK, 2010–2040, by attained age and projection scenario. Number of survivors (proportion of the population per 100,000).

		Year							
Age	Scenario	2010		2020		2030		2040	
0–44	1	23,023	(129)	28,037	(154)	38,653	(205)	43,816	(232)
	2	22,029	(124)	19,481	(107)	22,546	(120)	21,485	(114)
	3	23,029	(129)	27,390	(151)	36,850	(195)	41,055	(217)
	4	22,025	(124)	19,987	(110)	23,735	(126)	23,044	(122)
45–64	1	208,626	(2,588)	272,938	(3,188)	342,897	(4,116)	417,920	(4,960)
	2	208,338	(2,584)	236,170	(2,759)	238,501	(2,863)	232,085	(2,755)
	3	209,741	(2,602)	273,953	(3,200)	331,103	(3,974)	388,665	(4,613)
	4	207,224	(2,570)	235,271	(2,748)	247,194	(2,967)	250,074	(2,968)
≥65	1	338,234	(5,883)	539,484	(7,835)	830,769	(9,937)	1,221,001	(13,281)
	2	335,283	(5,832)	471,637	(6,850)	588,921	(7,044)	662,563	(7,207)
	3	335,958	(5,844)	497,549	(7,226)	696,934	(8,336)	915,927	(9,962)
	4	337,556	(5,872)	510,858	(7,420)	703,210	(8,411)	890,974	(9,691)
All	1	569,883	(1,803)	840,460	(2,500)	1,212,319	(3,406)	1,682,737	(4,598)
	2	565,649	(1,789)	727,288	(2,164)	849,967	(2,388)	916,133	(2,503)
	3	568,728	(1,799)	798,892	(2,377)	1,064,887	(2,992)	1,345,647	(3,677)
	4	566,804	(1,793)	766,116	(2,279)	974,139	(2,737)	1,164,093	(3,181)

	Inc.	Surv.	Popn.
Scenario 1:	↕	↕	↕
Scenario 2:	↔	↔	↕
Scenario 3:	↕	↔	↕
Scenario 4:	↔	↕	↕

Table 6.12. Rate of change of complete prevalence of female breast cancer in the UK, 2010–2040, by attained age and projection scenario. Average annual percentage change in number of survivors (proportion of the population).

		Period		
Age	Scenario	2010–2020	2020–2030	2030–2040
0–44	1	2.0 (1.8)	3.3 (2.9)	1.3 (1.2)
	2	-1.2 (-1.4)	1.5 (1.1)	-0.5 (-0.5)
	3	1.7 (1.5)	3.0 (2.6)	1.1 (1.1)
	4	-1.0 (-1.2)	1.7 (1.3)	-0.3 (-0.3)
45–64	1	2.7 (2.1)	2.3 (2.6)	2.0 (1.9)
	2	1.3 (0.7)	0.1 (0.4)	-0.3 (-0.4)
	3	2.7 (2.1)	1.9 (2.2)	1.6 (1.5)
	4	1.3 (0.7)	0.5 (0.8)	0.1 (0.0)
≥65	1	4.8 (2.9)	4.4 (2.4)	3.9 (2.9)
	2	3.5 (1.6)	2.2 (0.3)	1.2 (0.2)
	3	4.0 (2.1)	3.4 (1.4)	2.8 (1.8)
	4	4.2 (2.4)	3.2 (1.3)	2.4 (1.4)
All	1	4.0 (3.3)	3.7 (3.1)	3.3 (3.0)
	2	2.5 (1.9)	1.6 (1.0)	0.8 (0.5)
	3	3.5 (2.8)	2.9 (2.3)	2.4 (2.1)
	4	3.1 (2.4)	2.4 (1.8)	1.8 (1.5)

	Inc.	Surv.	Popn.
Scenario 1:	↕	↕	↕
Scenario 2:	↔	↔	↕
Scenario 3:	↕	↔	↕
Scenario 4:	↔	↕	↕

Table 6.13. Complete prevalence of all other malignant neoplasms (excluding non-melanoma skin cancer) in the UK, 2010–2040, by attained age, sex and projection scenario. Number of survivors (proportion of the population per 100,000).

		Year							
Age	Scenario	2010		2020		2030		2040	
Males									
0–44	1	66,849	(365)	70,634	(374)	79,430	(404)	80,786	(410)
	2	67,564	(369)	73,602	(390)	81,383	(414)	80,667	(410)
	3	67,337	(368)	73,038	(387)	82,075	(418)	83,001	(421)
	4	67,072	(367)	71,209	(377)	78,847	(401)	78,610	(399)
45–64	1	137,394	(1,765)	167,457	(2,033)	187,509	(2,289)	210,096	(2,487)
	2	138,585	(1,781)	169,458	(2,057)	179,390	(2,190)	185,316	(2,194)
	3	138,789	(1,783)	172,809	(2,098)	187,851	(2,293)	202,253	(2,395)
	4	137,199	(1,763)	164,064	(1,991)	178,434	(2,178)	191,396	(2,266)
≥65	1	224,437	(4,961)	340,718	(5,896)	495,003	(6,959)	675,515	(8,629)
	2	222,612	(4,921)	302,523	(5,235)	376,877	(5,299)	430,750	(5,502)
	3	224,432	(4,961)	319,633	(5,531)	416,193	(5,851)	497,992	(6,361)
	4	222,693	(4,922)	322,476	(5,581)	446,881	(6,283)	580,454	(7,415)
All	1	428,680	(1,401)	578,810	(1,759)	761,941	(2,178)	966,396	(2,684)
	2	428,760	(1,401)	545,583	(1,658)	637,649	(1,823)	696,732	(1,935)
	3	430,558	(1,407)	565,481	(1,718)	686,119	(1,961)	783,247	(2,175)
	4	426,965	(1,395)	557,749	(1,695)	704,162	(2,013)	850,460	(2,362)
Females									
0–44	1	67,712	(381)	71,769	(395)	81,221	(431)	82,748	(438)
	2	68,367	(384)	74,368	(410)	81,966	(435)	80,385	(425)
	3	68,112	(383)	73,475	(405)	82,777	(439)	83,672	(443)
	4	67,965	(382)	72,578	(400)	80,345	(426)	79,350	(420)
45–64	1	162,811	(2,019)	198,429	(2,318)	223,348	(2,681)	245,298	(2,911)
	2	163,540	(2,029)	194,527	(2,272)	206,441	(2,478)	209,782	(2,490)
	3	164,310	(2,038)	204,215	(2,385)	224,347	(2,693)	237,988	(2,825)
	4	162,070	(2,010)	189,056	(2,208)	205,090	(2,462)	215,287	(2,555)
≥65	1	286,214	(4,979)	401,872	(5,837)	561,701	(6,719)	763,514	(8,305)
	2	285,868	(4,973)	374,719	(5,442)	457,417	(5,471)	523,443	(5,693)
	3	285,939	(4,974)	383,355	(5,568)	490,441	(5,866)	591,804	(6,437)
	4	286,168	(4,978)	393,051	(5,709)	523,953	(6,267)	674,372	(7,335)
All	1	516,737	(1,635)	672,070	(1,999)	866,271	(2,434)	1,091,560	(2,983)
	2	517,775	(1,638)	643,614	(1,915)	745,824	(2,095)	813,610	(2,223)
	3	518,360	(1,640)	661,045	(1,967)	797,565	(2,241)	913,463	(2,496)
	4	516,204	(1,633)	654,685	(1,948)	809,389	(2,274)	969,010	(2,648)

	Inc.	Surv.	Popn.
Scenario 1:	↕	↕	↕
Scenario 2:	↔	↔	↕
Scenario 3:	↕	↔	↕
Scenario 4:	↔	↕	↕

Table 6.14. Rate of change of complete prevalence of all other malignant neoplasms (excluding non-melanoma skin cancer) in the UK, 2010–2040, by attained age, sex and projection scenario. Average annual percentage change in number of survivors (proportion of the population).

		Period		
Age	Scenario	2010–2020	2020–2030	2030–2040
Males				
0–44	1	0.6 (0.2)	1.2 (0.8)	0.2 (0.1)
	2	0.9 (0.5)	1.0 (0.6)	-0.1 (-0.1)
	3	0.8 (0.5)	1.2 (0.8)	0.1 (0.1)
	4	0.6 (0.3)	1.0 (0.6)	0.0 (0.0)
45–64	1	2.0 (1.4)	1.1 (1.2)	1.1 (0.8)
	2	2.0 (1.5)	0.6 (0.6)	0.3 (0.0)
	3	2.2 (1.6)	0.8 (0.9)	0.7 (0.4)
	4	1.8 (1.2)	0.8 (0.9)	0.7 (0.4)
≥65	1	4.3 (1.7)	3.8 (1.7)	3.2 (2.2)
	2	3.1 (0.6)	2.2 (0.1)	1.3 (0.4)
	3	3.6 (1.1)	2.7 (0.6)	1.8 (0.8)
	4	3.8 (1.3)	3.3 (1.2)	2.6 (1.7)
All	1	3.0 (2.3)	2.8 (2.2)	2.4 (2.1)
	2	2.4 (1.7)	1.6 (1.0)	0.9 (0.6)
	3	2.8 (2.0)	2.0 (1.3)	1.3 (1.0)
	4	2.7 (2.0)	2.4 (1.7)	1.9 (1.6)
Females				
0–44	1	0.6 (0.4)	1.2 (0.9)	0.2 (0.2)
	2	0.8 (0.6)	1.0 (0.6)	-0.2 (-0.2)
	3	0.8 (0.6)	1.2 (0.8)	0.1 (0.1)
	4	0.7 (0.5)	1.0 (0.6)	-0.1 (-0.1)
45–64	1	2.0 (1.4)	1.2 (1.5)	0.9 (0.8)
	2	1.8 (1.1)	0.6 (0.9)	0.2 (0.0)
	3	2.2 (1.6)	0.9 (1.2)	0.6 (0.5)
	4	1.6 (0.9)	0.8 (1.1)	0.5 (0.4)
≥65	1	3.5 (1.6)	3.4 (1.4)	3.1 (2.1)
	2	2.7 (0.9)	2.0 (0.1)	1.4 (0.4)
	3	3.0 (1.1)	2.5 (0.5)	1.9 (0.9)
	4	3.2 (1.4)	2.9 (0.9)	2.6 (1.6)
All	1	2.7 (2.0)	2.6 (2.0)	2.3 (2.1)
	2	2.2 (1.6)	1.5 (0.9)	0.9 (0.6)
	3	2.5 (1.8)	1.9 (1.3)	1.4 (1.1)
	4	2.4 (1.8)	2.1 (1.6)	1.8 (1.5)

	Inc.	Surv.	Popn.
Scenario 1:	↕	↕	↕
Scenario 2:	↔	↔	↕
Scenario 3:	↕	↔	↕
Scenario 4:	↔	↕	↕

Table 6.15. Complete prevalence of all malignant neoplasms (excluding non-melanoma skin cancer) in the UK, 2010–2040, by attained age, sex and projection scenario. Number of survivors (proportion of the population per 100,000).

		Year			
Age	Scenario	2010	2020	2030	2040
Males					
0–44	1	69,754 (381)	73,882 (391)	83,403 (424)	84,850 (431)
	2	70,531 (385)	77,289 (409)	85,800 (436)	85,038 (432)
	3	70,291 (384)	76,467 (405)	86,139 (438)	87,049 (442)
	4	69,989 (382)	74,658 (395)	83,060 (423)	82,861 (421)
45–64	1	207,686 (2,669)	250,189 (3,037)	283,366 (3,459)	306,771 (3,632)
	2	209,378 (2,690)	250,180 (3,037)	267,485 (3,265)	268,205 (3,175)
	3	210,125 (2,700)	256,735 (3,116)	280,308 (3,421)	291,155 (3,447)
	4	206,958 (2,659)	243,507 (2,956)	269,552 (3,290)	281,161 (3,329)
≥65	1	572,574 (12,656)	899,021 (15,558)	1,329,891 (18,698)	1,824,121 (23,301)
	2	570,780 (12,616)	805,986 (13,948)	1,007,612 (14,167)	1,154,177 (14,743)
	3	572,145 (12,647)	821,708 (14,220)	1,050,573 (14,771)	1,235,310 (15,780)
	4	571,211 (12,626)	881,320 (15,252)	1,275,170 (17,928)	1,705,835 (21,790)
All	1	850,014 (2,777)	1,223,093 (3,717)	1,696,659 (4,850)	2,215,742 (6,153)
	2	850,689 (2,779)	1,133,455 (3,444)	1,360,897 (3,890)	1,507,419 (4,186)
	3	852,561 (2,785)	1,154,910 (3,510)	1,417,020 (4,051)	1,613,514 (4,481)
	4	848,158 (2,771)	1,199,485 (3,645)	1,627,782 (4,653)	2,069,857 (5,748)
Females					
0–44	1	93,498 (525)	102,886 (567)	123,563 (655)	130,382 (690)
	2	93,092 (523)	96,840 (533)	107,858 (572)	105,100 (556)
	3	93,894 (528)	103,947 (573)	123,176 (653)	128,286 (679)
	4	92,689 (521)	95,552 (526)	107,543 (570)	105,833 (560)
45–64	1	399,260 (4,952)	506,357 (5,914)	608,052 (7,299)	709,393 (8,419)
	2	399,456 (4,955)	461,687 (5,393)	477,312 (5,729)	472,895 (5,613)
	3	402,082 (4,987)	511,853 (5,979)	592,612 (7,113)	664,714 (7,889)
	4	396,678 (4,920)	456,348 (5,330)	488,167 (5,860)	502,008 (5,958)
≥65	1	735,950 (12,801)	1,095,362 (15,909)	1,610,292 (19,261)	2,284,859 (24,852)
	2	732,353 (12,739)	984,492 (14,298)	1,213,387 (14,514)	1,376,301 (14,970)
	3	732,985 (12,750)	1,020,693 (14,824)	1,361,323 (16,283)	1,714,030 (18,643)
	4	735,358 (12,791)	1,056,114 (15,339)	1,435,729 (17,173)	1,840,060 (20,014)
All	1	1,228,708 (3,887)	1,704,606 (5,071)	2,341,907 (6,579)	3,124,634 (8,538)
	2	1,224,900 (3,875)	1,543,019 (4,590)	1,798,557 (5,053)	1,954,297 (5,340)
	3	1,228,961 (3,887)	1,636,493 (4,868)	2,077,111 (5,836)	2,507,030 (6,850)
	4	1,224,725 (3,874)	1,608,014 (4,784)	2,031,439 (5,707)	2,447,900 (6,689)

	Inc.	Surv.	Popn.
Scenario 1:	↕	↕	↕
Scenario 2:	↔	↔	↕
Scenario 3:	↕	↔	↕
Scenario 4:	↔	↕	↕

Table 6.16. Rate of change of complete prevalence of all malignant neoplasms (excluding non-melanoma skin cancer) in the UK, 2010–2040, by attained age, sex and projection scenario. Average annual percentage change in number of survivors (proportion of the population).

		Period		
Age	Scenario	2010–2020	2020–2030	2030–2040
Males				
0–44	1	0.6 (0.3)	1.2 (0.8)	0.2 (0.2)
	2	0.9 (0.6)	1.1 (0.6)	-0.1 (-0.1)
	3	0.8 (0.5)	1.2 (0.8)	0.1 (0.1)
	4	0.6 (0.3)	1.1 (0.7)	0.0 (0.0)
45–64	1	1.9 (1.3)	1.3 (1.3)	0.8 (0.5)
	2	1.8 (1.2)	0.7 (0.7)	0.0 (-0.3)
	3	2.0 (1.4)	0.9 (0.9)	0.4 (0.1)
	4	1.6 (1.1)	1.0 (1.1)	0.4 (0.1)
≥65	1	4.6 (2.1)	4.0 (1.9)	3.2 (2.2)
	2	3.5 (1.0)	2.3 (0.2)	1.4 (0.4)
	3	3.7 (1.2)	2.5 (0.4)	1.6 (0.7)
	4	4.4 (1.9)	3.8 (1.6)	3.0 (2.0)
All	1	3.7 (3.0)	3.3 (2.7)	2.7 (2.4)
	2	2.9 (2.2)	1.8 (1.2)	1.0 (0.7)
	3	3.1 (2.3)	2.1 (1.4)	1.3 (1.0)
	4	3.5 (2.8)	3.1 (2.5)	2.4 (2.1)
Females				
0–44	1	1.0 (0.8)	1.8 (1.5)	0.5 (0.5)
	2	0.4 (0.2)	1.1 (0.7)	-0.3 (-0.3)
	3	1.0 (0.8)	1.7 (1.3)	0.4 (0.4)
	4	0.3 (0.1)	1.2 (0.8)	-0.2 (-0.2)
45–64	1	2.4 (1.8)	1.8 (2.1)	1.6 (1.4)
	2	1.5 (0.9)	0.3 (0.6)	-0.1 (-0.2)
	3	2.4 (1.8)	1.5 (1.8)	1.2 (1.0)
	4	1.4 (0.8)	0.7 (1.0)	0.3 (0.2)
≥65	1	4.1 (2.2)	3.9 (1.9)	3.6 (2.6)
	2	3.0 (1.2)	2.1 (0.1)	1.3 (0.3)
	3	3.4 (1.5)	2.9 (0.9)	2.3 (1.4)
	4	3.7 (1.8)	3.1 (1.1)	2.5 (1.5)
All	1	3.3 (2.7)	3.2 (2.6)	2.9 (2.6)
	2	2.3 (1.7)	1.5 (1.0)	0.8 (0.6)
	3	2.9 (2.3)	2.4 (1.8)	1.9 (1.6)
	4	2.8 (2.1)	2.4 (1.8)	1.9 (1.6)

	Inc.	Surv.	Popn.
Scenario 1:	↕	↕	↕
Scenario 2:	↔	↔	↕
Scenario 3:	↕	↔	↕
Scenario 4:	↔	↕	↕

Table 6.17. Prevalence of colon, rectum and anus cancer in the UK, 2010–2040, by time since diagnosis, sex and projection scenario. Number of survivors (proportion of the population per 100,000).

Years since diagnosis	Scenario	Year 2010	2020	2030	2040
Males					
<1	1	16,788 (55)	23,105 (70)	30,811 (88)	37,671 (105)
	2	16,411 (54)	19,785 (60)	23,335 (67)	25,097 (70)
	3	16,758 (55)	21,770 (66)	27,554 (79)	32,324 (90)
	4	16,420 (54)	20,954 (64)	26,010 (74)	29,106 (81)
1–5	1	40,497 (132)	58,746 (179)	82,668 (236)	107,168 (298)
	2	40,595 (133)	50,038 (152)	58,238 (166)	63,358 (176)
	3	40,595 (133)	53,721 (163)	67,237 (192)	78,780 (219)
	4	40,497 (132)	54,465 (166)	71,123 (203)	85,412 (237)
≥5	1	69,791 (228)	106,503 (324)	160,486 (459)	232,361 (645)
	2	69,336 (227)	95,305 (290)	117,212 (335)	132,183 (367)
	3	69,340 (227)	96,718 (294)	125,214 (358)	150,665 (418)
	4	69,786 (228)	104,642 (318)	149,375 (427)	202,728 (563)
Females					
<1	1	12,635 (40)	15,526 (46)	19,236 (54)	22,146 (61)
	2	12,749 (40)	14,623 (44)	16,949 (48)	18,227 (50)
	3	12,534 (40)	14,549 (43)	17,094 (48)	18,861 (52)
	4	12,848 (41)	15,580 (46)	19,028 (53)	21,284 (58)
1–5	1	32,612 (103)	40,827 (121)	52,834 (148)	64,151 (175)
	2	32,801 (104)	38,297 (114)	44,048 (124)	47,805 (131)
	3	32,801 (104)	37,736 (112)	43,939 (123)	48,396 (132)
	4	32,612 (103)	41,270 (123)	52,632 (148)	62,712 (171)
≥5	1	71,193 (225)	95,246 (283)	127,640 (359)	168,659 (461)
	2	70,682 (224)	88,679 (264)	106,377 (299)	119,748 (327)
	3	70,678 (224)	87,715 (261)	104,747 (294)	118,548 (324)
	4	71,197 (225)	96,134 (286)	129,351 (363)	169,813 (464)

	Inc.	Surv.	Popn.
Scenario 1:	↕	↕	↕
Scenario 2:	↔	↔	↕
Scenario 3:	↕	↔	↕
Scenario 4:	↔	↕	↕

Table 6.18. Rate of change of prevalence of colon, rectum and anus cancer in the UK, 2010–2040, by time since diagnosis, sex and projection scenario. Average annual percentage change in number of survivors (proportion of the population).

Years since diagnosis	Scenario	Period 2010–2020	Period 2020–2030	Period 2030–2040
Males				
<1	1	3.2 (2.5)	2.9 (2.3)	2.0 (1.7)
	2	1.9 (1.2)	1.7 (1.0)	0.7 (0.4)
	3	2.7 (1.9)	2.4 (1.8)	1.6 (1.3)
	4	2.5 (1.7)	2.2 (1.6)	1.1 (0.8)
1–5	1	3.8 (3.0)	3.5 (2.8)	2.6 (2.3)
	2	2.1 (1.4)	1.5 (0.9)	0.8 (0.6)
	3	2.8 (2.1)	2.3 (1.6)	1.6 (1.3)
	4	3.0 (2.3)	2.7 (2.1)	1.8 (1.6)
≥5	1	4.3 (3.6)	4.2 (3.6)	3.8 (3.5)
	2	3.2 (2.5)	2.1 (1.5)	1.2 (0.9)
	3	3.4 (2.6)	2.6 (2.0)	1.9 (1.6)
	4	4.1 (3.4)	3.6 (3.0)	3.1 (2.8)
Females				
<1	1	2.1 (1.5)	2.2 (1.6)	1.4 (1.1)
	2	1.4 (0.8)	1.5 (0.9)	0.7 (0.4)
	3	1.5 (0.9)	1.6 (1.0)	1.0 (0.7)
	4	1.9 (1.3)	2.0 (1.4)	1.1 (0.8)
1–5	1	2.3 (1.6)	2.6 (2.0)	2.0 (1.7)
	2	1.6 (0.9)	1.4 (0.8)	0.8 (0.5)
	3	1.4 (0.8)	1.5 (1.0)	1.0 (0.7)
	4	2.4 (1.8)	2.5 (1.9)	1.8 (1.5)
≥5	1	3.0 (2.3)	3.0 (2.4)	2.8 (2.5)
	2	2.3 (1.7)	1.8 (1.3)	1.2 (0.9)
	3	2.2 (1.6)	1.8 (1.2)	1.2 (1.0)
	4	3.0 (2.4)	3.0 (2.4)	2.8 (2.5)

	Inc.	Surv.	Popn.
Scenario 1:	↕	↕	↕
Scenario 2:	↔	↔	↕
Scenario 3:	↕	↔	↕
Scenario 4:	↔	↕	↕

Table 6.19. Prevalence of lung, bronchus and trachea cancer in the UK, 2010–2040, by time since diagnosis, sex and projection scenario. Number of survivors (proportion of the population per 100,000).

Years since diagnosis	Scenario	Year 2010		2020		2030		2040	
Males									
<1	1	9,235	(30)	9,613	(29)	9,628	(28)	8,941	(25)
	2	9,741	(32)	11,869	(36)	14,065	(40)	15,183	(42)
	3	9,592	(31)	9,105	(28)	8,372	(24)	7,176	(20)
	4	9,349	(31)	12,528	(38)	16,211	(46)	19,034	(53)
1–5	1	7,967	(26)	8,289	(25)	8,550	(24)	8,292	(23)
	2	8,139	(27)	9,786	(30)	11,295	(32)	12,167	(34)
	3	8,139	(27)	8,173	(25)	7,368	(21)	6,255	(17)
	4	7,967	(26)	9,912	(30)	13,170	(38)	16,360	(45)
≥5	1	21,625	(71)	21,922	(67)	23,030	(66)	24,380	(68)
	2	21,669	(71)	20,453	(62)	21,822	(62)	24,132	(67)
	3	21,541	(70)	19,308	(59)	17,873	(51)	16,500	(46)
	4	21,743	(71)	23,131	(70)	28,186	(81)	36,224	(101)
Females									
<1	1	8,131	(26)	12,083	(36)	17,530	(49)	23,720	(65)
	2	7,727	(24)	8,868	(26)	10,158	(29)	10,769	(29)
	3	8,343	(26)	11,103	(33)	14,573	(41)	17,973	(49)
	4	7,541	(24)	9,706	(29)	12,353	(35)	14,402	(39)
1–5	1	7,074	(22)	11,621	(35)	18,075	(51)	26,387	(72)
	2	7,178	(23)	8,587	(26)	9,625	(27)	10,230	(28)
	3	7,178	(23)	10,957	(33)	14,127	(40)	17,399	(48)
	4	7,074	(22)	9,233	(27)	12,592	(35)	15,991	(44)
≥5	1	10,444	(33)	16,774	(50)	28,002	(79)	45,274	(124)
	2	10,339	(33)	13,064	(39)	15,609	(44)	17,775	(49)
	3	10,338	(33)	14,495	(43)	20,179	(57)	26,743	(73)
	4	10,446	(33)	15,290	(45)	21,957	(62)	30,595	(84)

	Inc.	Surv.	Popn.
Scenario 1:	↕	↕	↕
Scenario 2:	↔	↔	↕
Scenario 3:	↕	↔	↕
Scenario 4:	↔	↕	↕

Table 6.20. Rate of change of prevalence of lung, bronchus and trachea cancer in the UK, 2010–2040, by time since diagnosis, sex and projection scenario. Average annual percentage change in number of survivors (proportion of the population).

Years since diagnosis	Scenario	Period 2010–2020	Period 2020–2030	Period 2030–2040
Males				
<1	1	0.4 (-0.3)	0.0 (-0.6)	-0.7 (-1.0)
	2	2.0 (1.3)	1.7 (1.1)	0.8 (0.5)
	3	-0.5 (-1.2)	-0.8 (-1.4)	-1.5 (-1.8)
	4	3.0 (2.2)	2.6 (2.0)	1.6 (1.3)
1–5	1	0.4 (-0.3)	0.3 (-0.3)	-0.3 (-0.6)
	2	1.9 (1.1)	1.4 (0.8)	0.7 (0.5)
	3	0.0 (-0.7)	-1.0 (-1.6)	-1.6 (-1.9)
	4	2.2 (1.5)	2.9 (2.3)	2.2 (1.9)
≥5	1	0.1 (-0.6)	0.5 (-0.1)	0.6 (0.3)
	2	-0.6 (-1.3)	0.6 (0.0)	1.0 (0.7)
	3	-1.1 (-1.8)	-0.8 (-1.4)	-0.8 (-1.1)
	4	0.6 (-0.1)	2.0 (1.4)	2.5 (2.2)
Females				
<1	1	4.0 (3.4)	3.8 (3.2)	3.1 (2.8)
	2	1.4 (0.8)	1.4 (0.8)	0.6 (0.3)
	3	2.9 (2.3)	2.8 (2.2)	2.1 (1.8)
	4	2.6 (1.9)	2.4 (1.9)	1.5 (1.3)
1–5	1	5.1 (4.4)	4.5 (3.9)	3.9 (3.6)
	2	1.8 (1.2)	1.1 (0.6)	0.6 (0.3)
	3	4.3 (3.7)	2.6 (2.0)	2.1 (1.8)
	4	2.7 (2.1)	3.2 (2.6)	2.4 (2.1)
≥5	1	4.9 (4.2)	5.3 (4.7)	4.9 (4.6)
	2	2.4 (1.7)	1.8 (1.2)	1.3 (1.0)
	3	3.4 (2.8)	3.4 (2.8)	2.9 (2.6)
	4	3.9 (3.2)	3.7 (3.1)	3.4 (3.1)

	Inc.	Surv.	Popn.
Scenario 1:	↕	↕	↕
Scenario 2:	↔	↔	↕
Scenario 3:	↕	↔	↕
Scenario 4:	↔	↕	↕

Table 6.21. Prevalence of prostate cancer in the UK, 2010–2040, by time since diagnosis and projection scenario*. Number of survivors (proportion of the population per 100,000).

Years since diagnosis	Scenario	Year							
		2010		2020		2030		2040	
<1	1	33,373	(109)	41,680	(127)	50,228	(144)	54,925	(153)
	2	33,377	(109)	40,597	(123)	47,810	(137)	51,455	(143)
	3								
	4								
1–5	1	106,161	(347)	133,296	(405)	169,580	(485)	197,181	(548)
	2	107,436	(351)	126,851	(385)	148,294	(424)	161,248	(448)
	3								
	4								
≥5	1	115,898	(379)	241,128	(733)	399,736	(1,143)	578,427	(1,606)
	2	115,226	(376)	213,187	(648)	281,179	(804)	325,865	(905)
	3								
	4								

*Scenarios 3 and 4 are omitted for prostate cancer due to being the same as scenarios 2 and 1, respectively; see Table 6.1 for details.

	Inc.	Surv.	Popn.
Scenario 1:	↔	↕	↕
Scenario 2:	↔	↔	↕

Table 6.22. Rate of change of prevalence of prostate cancer in the UK, 2010–2040, by time since diagnosis and projection scenario*. Average annual percentage change in number of survivors (proportion of the population).

Years since diagnosis	Scenario	Period		
		2010–2020	2020–2030	2030–2040
<1	1	2.2 (1.5)	1.9 (1.3)	0.9 (0.6)
	2	2.0 (1.2)	1.6 (1.0)	0.7 (0.4)
	3			
	4			
1–5	1	2.3 (1.6)	2.4 (1.8)	1.5 (1.2)
	2	1.7 (0.9)	1.6 (1.0)	0.8 (0.5)
	3			
	4			
≥5	1	7.6 (6.8)	5.2 (4.5)	3.8 (3.5)
	2	6.3 (5.6)	2.8 (2.2)	1.5 (1.2)
	3			
	4			

*Scenarios 3 and 4 are omitted for prostate cancer due to being the same as scenarios 2 and 1, respectively; see Table 6.1 for details.

	Inc.	Surv.	Popn.
Scenario 1:	↔	↕	↕
Scenario 2:	↔	↔	↕

Table 6.23. Prevalence of female breast cancer in the UK, 2010–2040, by time since diagnosis and projection scenario. Number of survivors (proportion of the population per 100,000).

Years since diagnosis	Scenario	Year							
		2010		2020		2030		2040	
<1	1	45,892	(145)	60,881	(181)	80,136	(225)	101,636	(278)
	2	42,874	(136)	47,386	(141)	52,293	(147)	54,907	(150)
	3	45,898	(145)	60,118	(179)	78,195	(220)	98,190	(268)
	4	42,869	(136)	47,999	(143)	53,607	(151)	56,842	(155)
1–5	1	143,834	(455)	201,689	(600)	271,911	(764)	354,410	(968)
	2	144,222	(456)	160,835	(478)	174,831	(491)	184,322	(504)
	3	144,222	(456)	195,669	(582)	252,011	(708)	314,908	(860)
	4	143,834	(455)	165,720	(493)	188,652	(530)	207,460	(567)
≥5	1	380,157	(1,203)	577,889	(1,719)	860,271	(2,417)	1,226,691	(3,352)
	2	378,553	(1,197)	519,067	(1,544)	622,843	(1,750)	676,903	(1,850)
	3	378,609	(1,198)	543,105	(1,616)	734,682	(2,064)	932,550	(2,548)
	4	380,101	(1,202)	552,397	(1,643)	731,879	(2,056)	899,791	(2,459)

	Inc.	Surv.	Popn.
Scenario 1:	↕	↕	↕
Scenario 2:	↔	↔	↕
Scenario 3:	↕	↔	↕
Scenario 4:	↔	↕	↕

Table 6.24. Rate of change of prevalence of female breast cancer in the UK, 2010–2040, by time since diagnosis and projection scenario. Average annual percentage change in number of survivors (proportion of the population).

Years since diagnosis	Scenario	Period		
		2010–2020	2020–2030	2030–2040
<1	1	2.9 (2.2)	2.8 (2.2)	2.4 (2.1)
	2	1.0 (0.4)	1.0 (0.4)	0.5 (0.2)
	3	2.7 (2.1)	2.7 (2.1)	2.3 (2.0)
	4	1.1 (0.5)	1.1 (0.5)	0.6 (0.3)
1–5	1	3.4 (2.8)	3.0 (2.4)	2.7 (2.4)
	2	1.1 (0.5)	0.8 (0.3)	0.5 (0.3)
	3	3.1 (2.5)	2.6 (2.0)	2.3 (2.0)
	4	1.4 (0.8)	1.3 (0.7)	1.0 (0.7)
≥5	1	4.3 (3.6)	4.1 (3.5)	3.6 (3.3)
	2	3.2 (2.6)	1.8 (1.3)	0.8 (0.6)
	3	3.7 (3.0)	3.1 (2.5)	2.4 (2.1)
	4	3.8 (3.2)	2.9 (2.3)	2.1 (1.8)

	Inc.	Surv.	Popn.
Scenario 1:	↕	↕	↕
Scenario 2:	↔	↔	↕
Scenario 3:	↕	↔	↕
Scenario 4:	↔	↕	↕

Table 6.25. Prevalence of all other malignant neoplasms (excluding non-melanoma skin cancer) in the UK, 2010–2040, by time since diagnosis, sex and projection scenario. Number of survivors (proportion of the population per 100,000).

Years since diagnosis	Scenario	Year							
		2010		2020		2030		2040	
Males									
<1	1	49,273	(161)	65,614	(199)	85,185	(244)	103,028	(286)
	2	49,295	(161)	58,004	(176)	66,612	(190)	71,266	(198)
	3	50,982	(167)	63,234	(192)	76,660	(219)	87,449	(243)
	4	47,664	(156)	60,185	(183)	73,985	(211)	83,822	(233)
1–5	1	110,838	(362)	148,274	(451)	196,500	(562)	246,453	(684)
	2	111,215	(363)	132,805	(404)	148,969	(426)	159,545	(443)
	3	111,215	(363)	142,058	(432)	167,922	(480)	189,618	(527)
	4	110,838	(362)	138,390	(421)	173,615	(496)	205,983	(572)
≥5	1	268,569	(877)	364,921	(1,109)	480,256	(1,373)	616,916	(1,713)
	2	268,250	(876)	354,774	(1,078)	422,068	(1,207)	465,921	(1,294)
	3	268,361	(877)	360,189	(1,095)	441,537	(1,262)	506,179	(1,406)
	4	268,463	(877)	359,174	(1,091)	456,562	(1,305)	560,654	(1,557)
Females									
<1	1	46,388	(147)	58,420	(174)	73,429	(206)	87,675	(240)
	2	47,656	(151)	53,540	(159)	59,827	(168)	63,312	(173)
	3	48,208	(152)	57,278	(170)	67,792	(190)	76,584	(209)
	4	45,884	(145)	54,657	(163)	64,869	(182)	72,474	(198)
1–5	1	118,127	(374)	146,461	(436)	187,401	(526)	230,531	(630)
	2	118,750	(376)	136,611	(406)	149,903	(421)	157,964	(432)
	3	118,750	(376)	144,544	(430)	167,658	(471)	187,036	(511)
	4	118,127	(374)	138,499	(412)	167,551	(471)	194,459	(531)
≥5	1	352,222	(1,114)	467,189	(1,390)	605,441	(1,701)	773,354	(2,113)
	2	351,369	(1,111)	453,463	(1,349)	536,094	(1,506)	592,334	(1,619)
	3	351,402	(1,112)	459,223	(1,366)	562,114	(1,579)	649,844	(1,776)
	4	352,193	(1,114)	461,530	(1,373)	576,969	(1,621)	702,076	(1,918)

	Inc.	Surv.	Popn.
Scenario 1:	↕	↕	↕
Scenario 2:	↔	↔	↕
Scenario 3:	↕	↔	↕
Scenario 4:	↔	↕	↕

Table 6.26. Rate of change of prevalence of all other malignant neoplasms (excluding non-melanoma skin cancer) in the UK, 2010–2040, by time since diagnosis, sex and projection scenario. Average annual percentage change in number of survivors (proportion of the population).

Years since diagnosis	Scenario	Period 2010–2020	Period 2020–2030	Period 2030–2040
Males				
<1	1	2.9 (2.2)	2.6 (2.0)	1.9 (1.6)
	2	1.6 (0.9)	1.4 (0.8)	0.7 (0.4)
	3	2.2 (1.4)	1.9 (1.3)	1.3 (1.0)
	4	2.4 (1.6)	2.1 (1.5)	1.3 (1.0)
1–5	1	3.0 (2.2)	2.9 (2.2)	2.3 (2.0)
	2	1.8 (1.1)	1.2 (0.5)	0.7 (0.4)
	3	2.5 (1.7)	1.7 (1.1)	1.2 (0.9)
	4	2.2 (1.5)	2.3 (1.7)	1.7 (1.4)
≥5	1	3.1 (2.4)	2.8 (2.2)	2.5 (2.2)
	2	2.8 (2.1)	1.8 (1.1)	1.0 (0.7)
	3	3.0 (2.2)	2.1 (1.4)	1.4 (1.1)
	4	3.0 (2.2)	2.4 (1.8)	2.1 (1.8)
Females				
<1	1	2.3 (1.7)	2.3 (1.7)	1.8 (1.5)
	2	1.2 (0.6)	1.1 (0.5)	0.6 (0.3)
	3	1.7 (1.1)	1.7 (1.1)	1.2 (0.9)
	4	1.8 (1.1)	1.7 (1.1)	1.1 (0.8)
1–5	1	2.2 (1.5)	2.5 (1.9)	2.1 (1.8)
	2	1.4 (0.8)	0.9 (0.4)	0.5 (0.2)
	3	2.0 (1.4)	1.5 (0.9)	1.1 (0.8)
	4	1.6 (1.0)	1.9 (1.3)	1.5 (1.2)
≥5	1	2.9 (2.2)	2.6 (2.0)	2.5 (2.2)
	2	2.6 (2.0)	1.7 (1.1)	1.0 (0.7)
	3	2.7 (2.1)	2.0 (1.5)	1.5 (1.2)
	4	2.7 (2.1)	2.3 (1.7)	2.0 (1.7)

	Inc.	Surv.	Popn.
Scenario 1:	↕	↕	↕
Scenario 2:	↔	↔	↕
Scenario 3:	↕	↔	↕
Scenario 4:	↔	↕	↕

Table 6.27. Prevalence of all malignant neoplasms (excluding non-melanoma skin cancer) in the UK, 2010–2040, by time since diagnosis, sex and projection scenario. Number of survivors (proportion of the population per 100,000).

Years since diagnosis	Scenario	Year							
		2010		2020		2030		2040	
Males									
<1	1	108,669	(355)	140,013	(425)	175,851	(503)	204,564	(568)
	2	108,823	(356)	130,255	(396)	151,821	(434)	163,001	(453)
	3	110,708	(362)	134,705	(409)	160,395	(459)	178,404	(495)
	4	106,806	(349)	135,348	(411)	166,434	(476)	186,888	(519)
1–5	1	265,462	(867)	348,606	(1,059)	457,299	(1,307)	559,094	(1,553)
	2	267,385	(874)	319,480	(971)	366,796	(1,049)	396,318	(1,101)
	3	267,385	(874)	330,802	(1,005)	390,821	(1,117)	435,901	(1,211)
	4	265,462	(867)	336,062	(1,021)	427,489	(1,222)	504,936	(1,402)
≥5	1	475,884	(1,555)	734,474	(2,232)	1,063,509	(3,040)	1,452,083	(4,033)
	2	474,481	(1,550)	683,720	(2,078)	842,281	(2,408)	948,100	(2,633)
	3	474,469	(1,550)	689,403	(2,095)	865,804	(2,475)	999,208	(2,775)
	4	475,890	(1,555)	728,074	(2,213)	1,033,859	(2,955)	1,378,033	(3,827)
Females									
<1	1	113,046	(358)	146,910	(437)	190,331	(535)	235,178	(643)
	2	111,006	(351)	124,417	(370)	139,227	(391)	147,215	(402)
	3	114,983	(364)	143,049	(426)	177,654	(499)	211,608	(578)
	4	109,142	(345)	127,942	(381)	149,856	(421)	165,003	(451)
1–5	1	301,647	(954)	400,598	(1,192)	530,221	(1,490)	675,479	(1,846)
	2	302,951	(958)	344,329	(1,024)	378,407	(1,063)	400,321	(1,094)
	3	302,951	(958)	388,906	(1,157)	477,736	(1,342)	567,737	(1,551)
	4	301,647	(954)	354,721	(1,055)	421,427	(1,184)	480,623	(1,313)
≥5	1	814,015	(2,575)	1,157,097	(3,442)	1,621,355	(4,555)	2,213,978	(6,050)
	2	810,943	(2,565)	1,074,272	(3,196)	1,280,923	(3,599)	1,406,761	(3,844)
	3	811,027	(2,565)	1,104,538	(3,286)	1,421,722	(3,994)	1,727,684	(4,721)
	4	813,936	(2,575)	1,125,351	(3,348)	1,460,156	(4,102)	1,802,275	(4,925)

	Inc.	Surv.	Popn.
Scenario 1:	↕	↕	↕
Scenario 2:	↔	↔	↕
Scenario 3:	↕	↔	↕
Scenario 4:	↔	↕	↕

Table 6.28. Rate of change of prevalence of all malignant neoplasms (excluding non-melanoma skin cancer) in the UK, 2010–2040, by time since diagnosis, sex and projection scenario. Average annual percentage change in number of survivors (proportion of the population).

Years since diagnosis	Scenario	Period 2010–2020	Period 2020–2030	Period 2030–2040
Males				
<1	1	2.6 (1.8)	2.3 (1.7)	1.5 (1.2)
	2	1.8 (1.1)	1.5 (0.9)	0.7 (0.4)
	3	2.0 (1.2)	1.8 (1.1)	1.1 (0.8)
	4	2.4 (1.7)	2.1 (1.5)	1.2 (0.9)
1–5	1	2.8 (2.0)	2.8 (2.1)	2.0 (1.7)
	2	1.8 (1.1)	1.4 (0.8)	0.8 (0.5)
	3	2.2 (1.4)	1.7 (1.1)	1.1 (0.8)
	4	2.4 (1.6)	2.4 (1.8)	1.7 (1.4)
≥5	1	4.4 (3.7)	3.8 (3.1)	3.2 (2.9)
	2	3.7 (3.0)	2.1 (1.5)	1.2 (0.9)
	3	3.8 (3.1)	2.3 (1.7)	1.4 (1.2)
	4	4.3 (3.6)	3.6 (2.9)	2.9 (2.6)
Females				
<1	1	2.7 (2.0)	2.6 (2.0)	2.1 (1.9)
	2	1.1 (0.5)	1.1 (0.6)	0.6 (0.3)
	3	2.2 (1.6)	2.2 (1.6)	1.8 (1.5)
	4	1.6 (1.0)	1.6 (1.0)	1.0 (0.7)
1–5	1	2.9 (2.2)	2.8 (2.3)	2.5 (2.2)
	2	1.3 (0.7)	0.9 (0.4)	0.6 (0.3)
	3	2.5 (1.9)	2.1 (1.5)	1.7 (1.5)
	4	1.6 (1.0)	1.7 (1.2)	1.3 (1.0)
≥5	1	3.6 (2.9)	3.4 (2.8)	3.2 (2.9)
	2	2.9 (2.2)	1.8 (1.2)	0.9 (0.7)
	3	3.1 (2.5)	2.6 (2.0)	2.0 (1.7)
	4	3.3 (2.7)	2.6 (2.1)	2.1 (1.8)

	Inc.	Surv.	Popn.
Scenario 1:	↕	↕	↕
Scenario 2:	↔	↔	↕
Scenario 3:	↕	↔	↕
Scenario 4:	↔	↕	↕

6.5 Discussion

The results presented in this chapter provide a detailed set of projections of cancer prevalence in the UK for the next three decades, up to 2040. Four of the major cancer types in the UK – prostate, female breast, colorectal and lung – were considered separately, and a final category for all other cancers (excluding non-melanoma skin cancer) allowed estimates of prevalence to be produced for all malignant neoplasms combined. Prevalence counts and proportions (the number of survivors per 100,000 population) were calculated according to three attained age groups (0–44, 45–64 and ≥65 years) and three time since diagnosis bands (<1, 1–5 and ≥5 years), as well as by sex and cancer type.

These projections covered a 31 year period from 2009 to 2040 and, given the inherent uncertainty involved in such long-term forecasting, four different scenarios of future cancer incidence rates, cancer survival and population demographics – the three factors which directly influence cancer prevalence – were considered separately (Table 6.1). These scenarios allowed the effects on cancer prevalence of projected cancer incidence rates and survival to be examined separately to that of projected changes in population demographics.

In section 6.4, the results for scenario 1 were given special consideration. This was the scenario under which empirical trends in cancer incidence rates and survival were extrapolated, without attenuation, from 2009 up to 2040. (Prostate cancer incidence rates were kept static, since the extrapolation results were considered to be extremely unrealistic in this instance.) Projections of cancer prevalence under this scenario provided estimates based on the simplistic, and in places optimistic, assumption that existing trends in cancer incidence rates and survival will continue unabated for the next 30 years, and this should be kept in mind when considering the results. Nonetheless, this scenario is the most empirically motivated of those presented, and could therefore be considered to be the most likely.

Of course, it is impossible to anticipate precisely the future events and interventions that may effect changes in cancer prevalence – e.g. new screening programmes, public awareness campaigns or cancer treatments. If the prevalence estimates are viewed together with the results of the incidence and mortality regression models of section 5.5, then a greater understanding of the factors that influenced the scenario 1 projections can be gained. These differ for each cancer type, sex and age group, and so in certain instances alternative scenarios to projection scenario 1 may be considered

more likely based on other intelligence. For example, Mistry et al. (2011) estimated the separate effects on prostate cancer incidence of a) the introduction of the PSA test, and b) the underlying trend. They were then able to project prostate cancer incidence based on the explicit assumption that PSA testing levels would remain at current levels whilst the underlying trend would continue. A similar approach was used by these authors to project female breast cancer incidence taking into account the effect of breast cancer screening.

In the UK, cancer incidence rates are generally increasing. Of the cancers studied here, it was only male lung cancer incidence rates that exhibited a decreasing trend in the period 1974–2008 (Chapter 5: Figure 5.10). Cancer survival is also generally increasing, as can be seen from the decreasing mortality probabilities for survivors displayed in Chapter 5 (Figures 5.11 and 5.12). The reasons for these observed trends vary according to the cancer type in question. For example, recorded prostate cancer incidence rates have increased rapidly since the introduction of the PSA test as a screening tool and, since many cancers are now being diagnosed earlier, recorded survival has also increased (Cancer Research UK, 2010; Evans and Møller, 2003); recorded female breast cancer incidence rates have increased due to greater public awareness of the early symptoms and the introduction of a national screening programme in England (NHS Cancer Screening Programmes, 2011); and male lung cancer incidence rates have declined due, mainly, to a reduction in the prevalence of smoking among men in England since the 1970s (Davy, 2006). General increases in cancer survival have been brought about by advances in cancer treatment as well as a greater focus on earlier diagnosis. Each of these factors – increasing cancer incidence rates and increasing cancer survival – acts to increase cancer prevalence, since the former means that more cancers are being diagnosed and the latter that people are living longer with cancer.

Furthermore, the population of the UK is growing in size and also ageing (Office for National Statistics, 2011c). Combined, these demographic changes lead to increasing cancer prevalence since there are more people to be diagnosed with cancer and a greater proportion of these are in the older age groups for which cancer incidence rates are highest.

So trends in each of the three epidemiological factors that directly control cancer prevalence – incidence rates, survival and population demographics – are, in general in the UK, acting to increase it. For this reason, projection scenario 1 resulted in the highest estimates of cancer prevalence for almost all cancer types, sexes and age groups.

The main exception to this was male lung cancer, incidence rates of which have been decreasing in recent years. By considering the other three scenarios, the extents to which changes in cancer incidence rates, survival or population demographics are likely to affect cancer prevalence, if existing trends continue, can be quantified separately.

For example, consider the average annual percentage change in prevalence of all malignant neoplasms combined – Table 6.16. For females in all age groups combined, projection scenarios 3 (dynamic incidence rates and static survival) and 4 (static incidence rates and dynamic survival) resulted in almost identical rates of increase in the number of cancer survivors: 2.9% for scenario 3 and 2.8% for scenario 4 in the 2010s, 2.4% for both scenarios in the 2020s and 1.9% for both scenarios in the 2030s. This implies that, overall, the projected increase in female cancer prevalence due to increasing incidence rates is roughly the same as that due to increasing survival. By considering the projections of cancer prevalence under scenario 2 it can be seen that, even if incidence rates and survival were to remain constant from 2009 onwards, the number of survivors would still increase by between 0.8% and 2.3% per year. This is partly due to the increasing and ageing population. However, the prevalence *proportions* in each age group will also increase under scenario 2, implying that static incidence rates and survival by themselves would act to increase cancer prevalence. This illustrates that it is not necessary for incidence and survival to be increasing for cancer prevalence to be increasing – for example, if incidence and survival are both static in the future, but are at such levels that the inflow to the prevalent population is greater than the outflow, then cancer prevalence will increase; historical changes in incidence and survival would take many years to wash through before cancer prevalence could ever reach a steady state.

Whilst for all female cancers combined the effects on cancer prevalence of increasing incidence rates and survival are roughly of the same magnitude, this is not the case for every type of cancer. For example, the difference between the rate of change of male colorectal cancer prevalence under scenarios 2 (static incidence rates and survival) and 4 (static incidence rates and dynamic survival) is roughly twice that between scenarios 2 and 3 (dynamic incidence rates and static survival) (Table 6.6), with the number of survivors projected to increase to 221,000 in 2040 under scenario 2, 262,000 under scenario 3 and 317,000 under scenario 4 (Table 6.5). This implies that, if existing trends in cancer incidence and survival are to persist, then it will be higher cancer survival that is responsible for the largest proportion of the projected increases in male colorectal cancer prevalence. Similarly, prevalence of female colorectal cancer is

projected to be roughly the same under scenarios 2 and 3 since incidence rates are projected to be quite stable; it is therefore the increases in cancer survival which are responsible for the majority of the increases in female colorectal cancer prevalence as projected under scenario 1.

The age structure of the population of cancer survivors will, under projection scenario 1, become increasingly dominated by the oldest age groups: in 2040, 77% of all cancer survivors will be aged at least 65 years under this scenario. Perhaps even more notably, it is projected that in 2040 almost one quarter of all people in the UK aged at least 65 years will be cancer survivors – the equivalent figure for 2008 was one eighth (see Chapter 2). This result, in particular, highlights the potential for significant increases in the burden of cancer on health service and community care resources if current trends in cancer incidence and survival continue. It is vitally important, therefore, that careful plans are laid so that resources exist to meet the needs of cancer survivors in the future, particularly given the likely large increases in the number of survivors over (the current) retirement age and the impact of cancer on a person's fitness to work.

The precise needs of cancer survivors in the UK, and how best to meet them, is the subject of ongoing research (Richards et al., 2011) but still more needs to be done. The results presented in this chapter, and earlier in this thesis, provide some further insight. Time since diagnosis was shown to be a key indicator of the quantity of cancer related acute health care utilised by the population of cancer survivors in the UK (Chapter 3). The first year following diagnosis and the last year of life contained the highest levels of acute cancer related health service utilisation, but there was also a significant amount of usage in the period 1–5 years after diagnosis. Under projection scenario 1, the number of survivors in each of the time since diagnosis bands <1, 1–5 and ≥ 5 years will increase, but the number who are long-term survivors will increase at the fastest rate – by 2040, 69% of all survivors will be at least five years beyond diagnosis under this scenario, compared with 62% in 2009. This will have an impact on the types of health care required by cancer survivors, with a greater focus on rehabilitation and the long-term, post-treatment effects of cancer.

It is hoped that these projections will be of use to health service commissioners and resource planners but, as with any long-term epidemiological projections, there are limitations to this work. Assumptions regarding the likely future trends in cancer incidence rates and survival must be made before any projections of cancer prevalence are possible. The approach used in this work was to provide a range of estimates based on two different assumptions: a) that existing trends will persist in the future without

attenuation; and b) that existing trends will be arrested with incidence rates and survival remaining constant throughout the projected period. Scenario 1, based on the assumption that existing trends will persist, can be considered the 'most likely' of those presented, but at the same time, by considering the results for scenarios 2–4, the influences of projected incidence rates and survival can be examined independently. The most likely scenario may indeed change depending on the type of cancer, sex or age groups under consideration, or based on other intelligence. However, for much of this chapter the focus has remained on scenario 1.

There is an inconsistency in the definition of scenario 1 which assumed static incidence rates for prostate cancer at the same time as allowing survival to be dynamic, since trends in each are inter-linked. Recorded incidence rates of prostate cancer have been increasing in the UK, as has survival, but both of these trends are, to a large extent, a consequence of the introduction of the PSA test. This test has resulted in more cancers being diagnosed (leading to higher recorded incidence rates) but also in generally earlier diagnosis (leading to higher recorded survival due to a combination of earlier stage at diagnosis and lead time bias). Indeed, many cancers diagnosed following a PSA test may never have become otherwise symptomatic in the natural life of the patient (Barry, 2009; van Leeuwen et al., 2010). However, in order to provide plausible estimates of future prostate cancer prevalence, and by extension prevalence of all male cancers combined, it was necessary to exclude the projected prostate cancer incidence rates from scenario 1.

6.6 Summary

In this chapter, estimates of future cancer prevalence in the UK were presented in detail and discussed. A range of estimates were provided based on various different scenarios of future cancer incidence rates, survival and population demographics, and some consideration was given to defining the most likely of these scenarios. The extent to which each of the input quantities (future incidence rates, survival and population demographics) was responsible for the projected trends in cancer prevalence was also examined. These results highlighted the potential for substantial future increases in the prevalence of cancer in the UK and the increased burden this would bring to the health service.

This draws to a close the work on projections of cancer prevalence which spans Chapters 4–6, and also marks the end of the analysis section of this thesis. What

follows is the final chapter, in which the work as a whole is recapped and the main themes arising from it are highlighted.

Chapter 7. General discussion

In this final chapter, a brief recap of the main results is given together with a short discussion of their significance and the themes arising. The main strengths and limitations of the work are also highlighted.

7.1 Key findings

It was estimated that, at the end of 2008, there were approximately two million cancer survivors in the UK. Of these, around 800,000 were male and 1.2 million were female. Prostate and female breast cancers were the most prevalent types, accounting for 31% and 46% of male and female cancer survivors, respectively. Lung cancer was, with just 64,000 survivors, the least prevalent of the cancer types studied. Cancer prevalence varied greatly with attained age; less than 0.5% of people aged under 45 years were cancer survivors, but around 13% of those aged at least 65 were. The distribution of cancer survivors between time since diagnosis bands varied between cancer types and was largely a consequence of the survival characteristics particular to each.

To describe the population of cancer survivors in more detail, temporal phases of survivorship were defined based on time since diagnosis and time until death. It was shown that a large majority of cancer survivors (1.7 million of the 2.0 million in the UK) were, at the end of 2008, more than one year beyond diagnosis and more than one year from death. This phase of survivorship was characterised by little acute health service utilisation, the vast majority of which related to the first year after diagnosis or the last year before death. Survivors of cancer types with poorer prognoses experienced higher levels of acute health service utilisation soon after diagnosis but, after five years' survival, there was little difference in the proportion of time spent admitted to hospital for cancer related episodes of care, regardless of the type of cancer.

A model for projecting cancer prevalence was derived and applied using cancer registry data to project cancer prevalence in the UK up to the year 2040. Under the assumption that existing trends in cancer incidence rates and survival will continue, the number of cancer survivors was projected to increase by approximately one million per decade. By 2040, cancer prevalence was projected to be 5.3 million (6.2% of the male population and 8.5% of the female population). In addition to growing significantly in size, the population of cancer survivors was also projected to become older and, on average, further beyond diagnosis: 77% of cancer survivors will be at least 65 years old in 2040 under the assumption of continuing current trends in incidence and survival, compared

with 63% in 2009, and 69% will be at least five years beyond diagnosis, compared with 62% in 2009. Despite this, the *number* of cancer survivors less than five years beyond diagnosis was projected to more than double from almost 800,000 in 2009 to almost 1.7 million in 2040. Furthermore, under the same assumption, it was projected that in 2040 almost one quarter of all those aged over 65 years will be cancer survivors. Other results were also provided based on different assumptions regarding future trends in cancer incidence and survival. It was shown that, for females, the projected increases in cancer prevalence (all malignant neoplasms and all ages combined) were due equally to projected increases in incidence rates and projected increases in survival. In contrast, for males, projected increases in survival had a larger influence on projected prevalence than projected incidence rates. However, a variety of patterns was observed for different combinations of age group, sex and cancer type.

7.2 Strengths and limitations

The estimates of cancer prevalence presented in Chapter 2 are the most up-to-date for the UK and, notably, were produced from an analysis of a long time series of cancer registry data from all four constituent countries of the UK. The person-time analysis of linked cancer registry and hospital activity data from HES is, as far as the author is aware, the first of its type using this dataset and greatly enriched the basic prevalence estimates. Projections of cancer prevalence were produced in a flexible way, such that the effects of changing incidence rates, survival and population demographics could be assessed independently.

A limitation of this work is that the available hospital activity data only contained details of in-patient and day case episodes of care and no data were available for health service utilisation not involving a visit to hospital. A full description of cancer survivorship must consider other types of health and social care, as well as that received in hospital. The work in Chapter 3 should therefore be considered in conjunction with other recently published research, for example that concerning cancer survivors and primary health care (Khan et al., 2008, 2010, 2011a, b).

The projections of cancer prevalence in Chapter 6 were based on the projections of cancer incidence and mortality contained in Chapter 5. More complex modelling of each individual type of cancer could improve the projections of cancer prevalence by, for example, building in explicit assumptions regarding the impact of historical or likely future cancer screening programmes. However, there is a level of uncertainty in such

long-term projections that cannot be avoided entirely and, in this thesis, projections were produced for a variety of possible future scenarios.

7.3 Interpretation

Cancer survivorship is currently high on the public health agenda, and the need for a greater focus on the needs of cancer survivors has been championed by both statutory and voluntary organisations in the UK. Primary prevention of cancer remains an ultimate goal for society, but it is clear that many public health initiatives and advances in cancer treatments have resulted in higher cancer prevalence due to increased recorded incidence rates and survival. Accordingly, the characteristics of the population of cancer survivors have also changed. Many more people can expect to live further beyond their cancer diagnosis, and it is no longer the death sentence it was once perceived to be.

Despite this, there has been little systematic study of cancer prevalence in the UK. This thesis, by describing cancer prevalence in terms of demographics, temporal features and acute health service utilisation, adds much to the understanding of the cancer survivor population and survivorship in the UK today. It has answered partly or wholly some of the research questions (as set out by the NCSI) which have been identified as key to ensuring that those living with cancer in the UK experience the best possible health outcomes now and in the future.

Despite being a large population, the majority of cancer survivors are neither recently diagnosed nor in the final year of their life and experience very little cancer related acute health service care. Most of the cancer burden to the acute health service comes from initial treatment and end of life care, and as general life expectancy increases and prognoses for many cancer types improve, the proportion of cancer survivors who are in the last year of their life decreases. However, the extent to which cancer survivors may require, or are receiving, other forms of care was not considered in this work. In the medium and long term, cancer survivors face day-to-day struggles – such as relationship and financial difficulties, or problems returning to work – and addressing these falls to community and social care services.

Increasing cancer prevalence, and the changing characteristics of cancer survivorship, present challenges to health service providers and society in general. Despite evidence that only a small proportion of cancer survivors have high levels of acute health service utilisation, the projections of cancer prevalence contained in this thesis highlight the

potential for significantly elevated demands on health service resources in the future. The *number* of cancer survivors requiring initial and ongoing treatment during the first five years after diagnosis will increase substantially as the population of the UK grows and ages, and as cancer incidence rates and survival continue to increase in general. It is clear, therefore, that adequate planning to ensure that the best possible use is made of available health service resources for cancer survivors is essential. Furthermore, as the population of cancer survivors becomes older, cancer will increasingly need to be considered as one of multiple co-morbidities, and efficient integration of different clinical specialities will be vital.

7.4 Summary

This thesis contains a detailed description of cancer prevalence and selected aspects of cancer survivorship in the UK. Various methods, concepts and models were developed and could form the basis of future research. The results presented here not only provide valuable insights into the characteristics of the *current* population of cancer survivors, but also describe some of the most important factors that will influence future changes in these characteristics. Society faces many challenges related to the growing and ageing general population of the UK – a large growth in the number of cancer survivors is not least among them. The coming decades are likely to bring difficult financial circumstances which will, therefore, test the ability of statutory and voluntary organisations to meet the diverse needs of those diagnosed with cancer. The work contained in this thesis, however, provides valuable intelligence that will help society meet these challenges.

Bibliography

Barry MJ (2009). 'Screening for Prostate Cancer – The Controversy That Refuses to Die'. *New England Journal of Medicine*; 360(13): 1351–54.

Benhamiche-Bouvier AM, Clinard F, Phelip JM, Rassiat E and Faivre J (2000). 'Colorectal cancer prevalence in France'. *European Journal of Cancer Prevention*; 9(5): 303–07.

Brenner H and Gefeller O (1997). 'Deriving more up to date estimates of long-term patient survival'. *Journal of Clinical Epidemiology*; 50(2): 211–16.

Brenner H and Hakulinen T (2004). 'Are patients diagnosed with breast cancer before age 50 years ever cured?'. *Journal of Clinical Oncology*; 22(3): 432–38.

Brown ML, Riley GF, Potosky AL and Etzioni RD (1999). 'Obtaining long-term disease specific costs of care: application to Medicare enrollees diagnosed with colorectal cancer'. *Medical Care*; 37(12): 1249–59.

Budescu DV (1985). 'Analysis of Dichotomous Variables in the Presence of Serial Dependence'. *Psychological Bulletin*; 97(3): 547–61.

Cancer Research UK (2009a). 'Cancer survival rates – trends'. [Accessed 2 December, 2011]; http://info.cancerresearchuk.org/cancerstats/survival/fiveyear/.

Cancer Research UK (2009b). 'Lung cancer survival statistics'. [Accessed 30 November, 2011]; http://info.cancerresearchuk.org/cancerstats/types/lung/survival/.

Cancer Research UK (2010). 'Prostate cancer – survival statistics'. [Accessed 2 December, 2011]; http://info.cancerresearchuk.org/cancerstats/types/prostate/survival/.

Cancer Research UK (2011a). 'All cancers combined statistics – Key Facts'. [Accessed 22 November, 2011]; http://info.cancerresearchuk.org/cancerstats/keyfacts/Allcancerscombined/.

Cancer Research UK (2011b). 'Cancer incidence for all cancers combined'. [Accessed 22 November, 2011]; http://info.cancerresearchuk.org/cancerstats/incidence/all-cancers-combined/.

Cancer Research UK (2011c). 'Prostate cancer – UK incidence statistics'. [Accessed 2 December, 2011]; http://info.cancerresearchuk.org/cancerstats/types/prostate/incidence/.

Capocaccia R and De Angelis R (1997). 'Estimating the completeness of prevalence based on cancer registry data'. *Statistics in Medicine*; 16(4): 425–40.

Capocaccia R, De Angelis R, Frova L, Gatta G, Sant M, Micheli A, Berrino F, Conti E, Gafa L, Roncuccl L, et al. (1997). 'Estimation and projections of colorectal cancer trends in Italy'. *International Journal of Epidemiology*; 26(5): 924–32.

Capocaccia R, De Angelis R, Frova L, Sant M, Buiatti E, Gatta G, Micheli A, Berrino F, Barchielli A, Centi E, et al. (1995). 'Estimation and projections of stomach cancer trends in Italy'. *Cancer Causes and Control*; 6(4): 339–46.

Carstensen, B (2007). 'Examples and programs for splitting follow-up time in cohort studies'. [Accessed 2 December, 2011]; http://staff.pubhealth.ku.dk/~bxc/Lexis/.

Clayton D and Schifflers E (1987). 'Models for temporal variation in cancer rates. II: Age-period-cohort models'. *Statistics in Medicine*; 6: 469–81.

Coldman A, McBride ML and Braun T (1992). 'Calculating The Prevalence of Cancer'. *Statistics in Medicine*; 11: 1579–89.

Collett D (2003). *Modelling Survival Data in Medical Research.* (Second edn., Boca Raton: Chapman and Hall/CRC).

Colonna M, Danzon A, Delafosse P, Mitton N, Bara S, Bouvier AM, Ganry O, Guizard AV, Launoy G, Molinie F, et al. (2008). 'Cancer prevalence in France: Time trend, situation in 2002 and extrapolation to 2012'. *European Journal of Cancer*; 44(1): 115–22.

Colonna M, Grosclaude P, Launoy G, Tretarre B, Arveux P, Raverdy N, Benhamiche AM, Herbert C and Faivre J (2001). 'Estimation of colorectal cancer prevalence in France'. *European Journal of Cancer*; 37(1): 93–96.

Colonna M, Hedelin G, Esteve J, Grosclaude P, Launoy G, Buemi A, Arveux P, Tretarre B, Chaplain G, Lesec'h JM, et al. (2000). 'National cancer prevalence estimation in France'. *International Journal of Cancer*; 87(2): 301–04.

Colonna M, Mitton N, Schott AM, Remontet L, Olive F, Gomez F, Iwaz J, Polazzi S, Bossard N and Trombert B (2012). 'Joint use of epidemiological and hospital medico-

administrative data to estimate prevalence. Application to French data on breast cancer'. *Cancer Epidemiology*; 36(2): 116-21.

Cooley C (2010a). 'Cancer survivorship 1: How services need to change for those living with and beyond cancer'. *Nursing Times*; 106(20): 24–25.

Cooley C (2010b). 'Cancer survivorship 2: providing advice and support to those living with and beyond cancer'. *Nursing Times*; 106(21): 17–18.

Davy M (2006). 'Time and generational trends in smoking among men and women in Great Britain, 1972–2004/05'. *Office for National Statistics Health Statistics Quarterly*; 32: 35–43.

De Angelis G, De Angelis R, Frova L and Verdecchia A (1994). 'MIAMOD: A computer package to estimate chronic disease morbidity using mortality and survival data'. *Computer Methods and Programs in Biomedicine*; 44(2): 99–107.

De Angelis R, Tavilla A, Verdecchia A, Scoppa S, Hachey M, Feuer EJ and Mariotto AB (2009). 'Breast cancer survivors in the United States: geographic variability and time trends, 2005-2015'. *Cancer*; 115(9): 1954–66.

Department of Health (2007). 'Cancer Reform Strategy'. [Accessed 2 December, 2011]; http://www.dh.gov.uk/en/ Publicationsandstatistics/Publications/ PublicationsPolicyAndGuidance/DH_081006.

Department of Health (2010). 'National Cancer Survivorship Initiative Vision'. [Accessed 19 December, 2011]; http://www.ncsi.org.uk/wp-content/uploads/NCSI-Vision-Document.pdf.

Dubeau CE (2006). 'The aging lower urinary tract'. *Journal of Urology*; 175(3 Pt 2): S11–5.

Duffy SW, Nagtegaal ID, Wallis M, Cafferty FH, Houssami N, Warwick J, Allgood PC, Kearins O, Tappenden N, O'Sullivan E, et al. (2008). 'Correcting for lead time and length bias in estimating the effect of screen detection on cancer survival'. *American Journal of Epidemiology*; 168(1): 98–104.

Elgar F (1968). *Mondrian.* (London: Thames and Hudson).

Evans HS, Lewis CM, Robinson D, Bell CM, Møller H and Hodgson SV (2001). 'Incidence of multiple primary cancers in a cohort of women diagnosed with breast cancer in southeast England'. *British Journal of Cancer*; 84(3): 435–40.

Evans HS and Møller H (2003). 'Recent trends in prostate cancer incidence and mortality in southeast England'. *European Urology*; 43(4): 337–41.

Evans HS, Møller H, Robinson D, Lewis CM, Bell CM and Hodgson SV (2002). 'The risk of subsequent primary cancers after colorectal cancer in southeast England'. *Gut*; 50(5): 647–52.

Feller W (1971). *An Introduction to Probability Theory and Its Applications.* (Second edn., New York: John Wiley & Sons).

Ferlay J, Shin HR, Bray F, Forman D, Mathers C and Parkin DM (2010). 'Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008'. *International Journal of Cancer*; 127(12): 2893–917.

Fiorentino F, Maddams J, Møller H and Utley M (2011). 'Modelling to estimate future trends in cancer prevalence'. *Health Care Management Science*; 14(3): 262–66.

Fitzpatrick D, Gavin A and Donnelly D (2006). *Cancer Trends in Northern Ireland: 1993–2003.* (Belfast: Northern Ireland Cancer Registry).

Forman D, Stockton D, Møller H, Quinn M, Babb P, De Angelis R and Micheli A (2003). 'Cancer prevalence in the UK: Results from the EUROPREVAL study'. *Annals of Oncology*; 14(4): 648–54.

Foster C, Wright D, Hill H, Hopkinson J and Roffe L (2009). 'Psychosocial implications of living 5 years or more following a cancer diagnosis: a systematic review of the research evidence'. *European Journal of Cancer Care*; 18(3): 223–47.

Francisci S, Capocaccia R, Grande E, Santaquilani M, Simonetti A, Allemani C, Gatta G, Sant M, Zigon G, Bray F, et al. (2009). 'The cure of cancer: A European perspective'. *European Journal of Cancer*; 45(6): 1067–79.

Fritz A, Percy C, Jack A, Shanmugaratnam K, Sobin L, Parkin DM and Whelan S (2000). *International Classification of Diseases for Oncology.* (Third edn., Geneva: World Health Organisation).

Gallivan S, Utley M, Jit M and Pagel C (2007). 'A Computational Algorithm Associated with Patient Progress Modelling'. *Computational Management Science*; 4(3): 283–99.

Gatenby PA, Hainsworth A, Caygill C, Watson A and Winslet M (2011). 'Projections for oesophageal cancer incidence in England to 2033'. *European Journal of Cancer Prevention*; 20(4): 283–86.

Gatta G, Capocaccia R, Berrino F, Ruzza MR, Contiero P, Oberaigner W, Storm H, Engholm G, Aareleid T, Hakulinen T, et al. (2004). 'Colon cancer prevalence and estimation of differing care needs of colon cancer patients'. *Annals of Oncology*; 15(7): 1136–42.

Grande E, Inghelmann R, Francisci S, Verdecchia A, Micheli A, Capocaccia R and De Angelis R (2006). 'Estimating regional cancer burden in countries with partial registration coverage: An application to all malignant neoplasms in Italy over the period 1970–2010'. *European Journal of Cancer*; 42(18): 3236–45.

Gregurek R, Bras M, Dordevic V, Ratkovic AS and Brajkovic L (2010). 'Psychological problems of patients with cancer'. *Psychiatria Danubina*; 22(2): 227–30.

Griebling T (2008). 'Urinary incontinence and voiding dysfunction in elderly men'. *Current Bladder Dysfunction Reports*; 3(4): 241–46.

Heinavaara S and Hakulinen T (2006). 'Predicting the lung cancer burden: Accounting for selection of the patients with respect to general population mortality'. *Statistics in Medicine*; 25(17): 2967–80.

Hubbard G (2010). 'Cancer: no longer a death sentence'. *British Journal of Community Nursing*; 15(5): 213.

Information Services Division Scotland (2011). 'Inpatient, Day Case and Outpatient Activity'. [Accessed 2 December, 2011]; http://www.isdscotland.org/isd/3405.html.

Jackson RRP and Aspden P (1979). 'Treatment Evaluation – A Modelling Approach with Application to Acute Myeloid Leukaemia'. *Journal of the Operational Research Society*; 30: 11–22.

Jackson RRP, Birkhead BG, Bell R, Lister TA and Gregory WM (1982). 'Application of Jackson, Aspden acute myeloid leukaemia model'. *Journal of the Operational Research Society*; 33: 211–16.

Jackson RRP, McKay VL and Malpas JS (1981). 'Patient progress modelling and results in Wilm's tumour'. *Omega*; 9(5): 527–36.

Jenkins D, Sherlaw-Johnson C and Gallivan S (1994). 'Can additional detection of human Papillomavirus type 16 improve cervical cytological screening?'. *Journal of Pathology*; 172(Suppl.): 134.

Keiding N (2006). 'Event history analysis and the cross-section'. *Statistics in Medicine*; 25: 2343–64.

Khan NF, Carpenter L, Watson E and Rose PW (2010). 'Cancer screening and preventative care among long-term cancer survivors in the United Kingdom'. *British Journal of Cancer*; 102(7): 1085–90.

Khan NF, Evans J and Rose PW (2011a). 'A qualitative study of unmet needs and interactions with primary care among cancer survivors'. *British Journal of Cancer*; 105(Suppl.1): S46–51.

Khan NF, Mant D, Carpenter L, Forman D and Rose PW (2011b). 'Long-term health outcomes in a British cohort of breast, colorectal and prostate cancer survivors: a database study'. *British Journal of Cancer*; 105(Suppl.1): S29–37.

Khan NF, Ward A, Watson E, Austoker J and Rose PW (2008). 'Long-term survivors of adult cancers and uptake of primary health services: a systematic review'. *European Journal of Cancer*; 44(2): 195–204.

Klotz J (1973). 'Statistical Inference in Bernoulli Trials with Dependence'. *Annals of Statistics*; 1(2): 373–79.

Lagiou P and Adami H (2002). 'Burden of Cancer' in Adami H, Hunter D and Trichopoulos D (eds.), *Textbook of Cancer Epidemiology* (New York: Oxford University Press), 3–28.

Lambert PC, Dickman PW, Osterlund P, Andersson T, Sankila R and Glimelius B (2007). 'Temporal trends in the proportion cured for cancer of the colon and rectum: a population-based study using data from the Finnish Cancer Registry'. *International Journal of Cancer*; 121(9): 2052–59.

Last JM (2001). *A Dictionary of Epidemiology.* (Fourth edn., New York: Oxford University Press).

Lexis W (1875). *Einleitung in die Theorie der Bevölkerungsstatistik.* (Strassburg: Trübner).

Liu L, de Vries E, Louwman M, Aben K, Janssen-Heijnen M, Brink M, Coebergh JW and Soerjomataram I (2011). 'Prevalence of multiple malignancies in the Netherlands in 2007'. *International Journal of Cancer*; 128(7): 1659–67.

Louchini R, Beaupre M, Demers AA, Goggin P and Bouchard C (2006). 'Trends in cancer prevalence in Quebec'. *Chronic Diseases in Canada*; 27(3): 110–19.

Maddams J, Brewster D, Gavin A, Steward J, Elliott J, Utley M and Møller H (2009). 'Cancer prevalence in the United Kingdom: estimates for 2008'. *British Journal of Cancer*; 101(3): 541–47.

Maddams J, Utley M and Møller H (2011a). 'A person-time analysis of hospital activity among cancer survivors in England'. *British Journal of Cancer*; 105(Suppl.1): S38–45.

Maddams J, Utley M and Møller H (2011b). 'Levels of acute health service use among cancer survivors in the United Kingdom'. *European Journal of Cancer*; 47(14): 2211–20.

Maliski SL, Heilemann MV and McCorkle R (2002). 'From "death sentence" to "good cancer": couples' transformation of a prostate cancer diagnosis'. *Nursing Research*; 51(6): 391–97.

Mariotto A, Warren JL, Knopf KB and Feuer EJ (2003). 'The prevalence of patients with colorectal carcinoma under care in the US'. *Cancer*; 98(6): 1253–61.

Mariotto AB, Yabroff KR, Feuer EJ, De Angelis R and Brown M (2006). 'Projecting the number of patients with colorectal carcinoma by phases of care in the US: 2000–2020'. *Cancer Causes and Control*; 17(10): 1215–26.

Markman M (2006). 'Viewing ovarian cancer as a "chronic disease": What exactly does this mean?'. *Gynecologic Oncology*; 100(2): 229–30.

Mayfield HF (1961). 'Nesting Success Calculated from Exposure'. *The Wilson Bulletin*; 73(3): 255–61.

Mayfield HF (1975). 'Suggestions for Calculating Nest Success'. *The Wilson Bulletin*; 87(4): 456–66.

McCabe MS (2007). 'Living beyond cancer: survivorship is more than surviving'. *Nature Clinical Practice Urology*; 4(11): 575.

McCarthy M (2002). 'New measure of cancer prevalence may aid health planning'. *Lancet Oncology*; 3(2): 70.

Merrill RM (2001). 'Partitioned prostate cancer prevalence estimates: An informative measure of the disease burden'. *Journal of Epidemiology and Community Health*; 55(3): 191–97.

Micheli A, Mugno E, Krogh V, Quinn MJ, Coleman M, Hakulinen T, Gatta G, Berrino F and Capocaccia R (2002). 'Cancer prevalence in European registry areas'. *Annals of Oncology*; 13(6): 840–65.

Mistry M, Parkin DM, Ahmad AS and Sasieni P (2011). 'Cancer incidence in the United Kingdom: projections to the year 2030'. *British Journal of Cancer*; 105(11): 1795–803.

Mitry E, Rachet B, Quinn MJ, Cooper N and Coleman MP (2008a). 'Survival from cancer of the colon in England and Wales up to 2001'. *British Journal of Cancer*; 99(Suppl.1): S26–S29.

Mitry E, Rachet B, Quinn MJ, Cooper N and Coleman MP (2008b). 'Survival from cancer of the rectum in England and Wales up to 2001'. *British Journal of Cancer*; 99(Suppl.1): S30–S32.

Møller B, Fekjaer H, Hakulinen T, Tryggvadottir L, Storm HH, Talback M and Haldorsen T (2002). 'Prediction of cancer incidence in the Nordic countries up to the year 2020'. *European Journal of Cancer Prevention*; 11(Suppl.1): S1–96.

Möller T, Anderson H, Aareleid T, Hakulinen T, Storm H, Tryggvadottir L, Corazziari I and Mugno E (2003). 'Cancer prevalence in Northern Europe: the EUROPREVAL study'. *Annals of Oncology*; 14(6): 946–57.

Mudholkar GS and Srivastava DK (1993). 'Exponentiated Weibull family for analyzing bathtub failure-rate data'. *IEEE Transactions on Reliability*; 42(2): 299–302.

National Cancer Intelligence Network (2011a). 'National Cancer Data Repository'. [Accessed 13 October, 2011]; http://www.ncin.org.uk/collecting_and_using_data/national_cancer_data_repository/default.aspx.

National Cancer Intelligence Network (2011b). 'UK Cancer Information Service'. [Accessed 18 November, 2011]; http://www.ncin.org.uk/cancer_information_tools/ukcis.aspx.

National Cancer Survivorship Initiative (2011). 'How we discuss survivorship'. [Accessed 18 November, 2011]; http://www.ncsi.org.uk/communicating-the-ncsi-vision/.

NHS Cancer Screening Programmes (2011). 'NHS Breast Cancer Screening Programme'. [Accessed 25 November, 2011]; http://www.cancerscreening.nhs.uk/breastscreen/about-breast-screening.html.

NHS Health and Social Care Information Centre (2010). 'HES User Guide'. [Accessed 20 December, 2011]; http://www.hesonline.nhs.uk/Ease/servlet/ContentServer?siteID=1937&categoryID=87.

NHS Scotland (2004). *Cancer Incidence Projections for Scotland (2001–2020). An Aid to Planning Cancer Services.* (Edinburgh: Scottish Executive).

Norlund A, Alvegard T, Lithman T, Merlo J and Noreen D (2003). 'Prostate Cancer: Prevalence-Based Healthcare Costs'. *Scandinavian Journal of Urology and Nephrology*; 37(5): 371-75.

Nyatanga B (2010). 'Survivorship: living with and beyond cancer'. *International Journal of Palliative Nursing*; 16(2): 55.

Office for National Statistics (2011a). 'Population Estimates'. [Accessed 2 December, 2011]; http://www.statistics.gov.uk/hub/population/population-change/population-estimates/index.html.

Office for National Statistics (2011b). 'Population Projections'. [Accessed 2 December, 2011]; http://www.statistics.gov.uk/hub/population/population-change/population-projections/index.html.

Office for National Statistics (2011c). 'Theme: Population'. [Accessed 2 December, 2011]; http://www.statistics.gov.uk/hub/population/index.html.

Olsen AH, Parkin DM and Sasieni P (2008). 'Cancer mortality in the United Kingdom: projections to the year 2025'. *British Journal of Cancer*; 99(9): 1549–54.

Parker C (2004). 'Active surveillance: towards a new paradigm in the management of early prostate cancer'. *Lancet Oncology*; 5(2): 101–06.

Parkin DM, Bray F, Ferlay J and Pisani P (2001). 'Estimating the world cancer burden: Globocan 2000'. *International Journal of Cancer*; 94(2): 153–56.

Parkin DM, Tappenden P, Olsen AH, Patnick J and Sasieni P (2008). 'Predicting the impact of the screening programme for colorectal cancer in the UK'. *Journal of Medical Screening*; 15(4): 163–74.

Phillips JL and Currow DC (2010). 'Cancer as a chronic disease'. *Collegian: Journal of the Royal College of Nursing Australia*; 17(2): 47–50.

Phillips N, Coldman A and McBride ML (2002). 'Estimating cancer prevalence using mixture models for cancer survival'. *Statistics in Medicine*; 21(9): 1257–70.

Pisani P, Bray F and Parkin DM (2002). 'Estimates of the world-wide prevalence of cancer for 25 sites in the adult population'. *International Journal of Cancer*; 97(1): 72–81.

Quinn MJ, Cooper N, Rachet B, Mitry E and Coleman MP (2008). 'Survival from cancer of the breast in women in England and Wales up to 2001'. *British Journal of Cancer*; 99(Suppl.1): S53–S55.

Rachet B, Maringe C, Nur U, Quaresma M, Shah A, Woods LM, Ellis L, Walters S, Forman D, Steward J, et al. (2009). 'Population-based cancer survival trends in England and Wales up to 2007: an assessment of the NHS cancer plan for England'. *Lancet Oncology*; 10(4): 351–69.

Rachet B, Quinn MJ, Cooper N and Coleman MP (2008a). 'Survival from cancer of the lung in England and Wales up to 2001'. *British Journal of Cancer*; 99(Suppl.1): S40–S42.

Rachet B, Woods LM, Mitry E, Riga M, Cooper N, Quinn MJ, Steward J, Brenner H, Esteve J, Sullivan R, et al. (2008b). 'Cancer survival in England and Wales at the end of the 20th century'. *British Journal of Cancer*; 99(Suppl.1): S2–10.

Ramsey SD, Berry K and Etzioni R (2002). 'Lifetime cancer-attributable cost of care for long term survivors of colorectal cancer'. *American Journal of Gastroenterology*; 97(2): 440-45.

Rebbeck TR, Sellers TA and Paskett ED (eds.) (2011). 'CEBP Focus: Cancer Survivorship Research' in *Cancer Epidemiology, Biomarkers and Prevention;* 20(10): 1994–2104.

Richards M, Corner J and Maher J (eds.) (2011). 'The National Cancer Survivorship Initiative: New and Emerging Evidence on the Ongoing Needs of Cancer Survivors' in *British Journal of Cancer;* 105(Suppl.1): S1–S94.

Robertson C, Gandini S and Boyle P (1999). 'Age-period-cohort models: a comparative study of available methodologies'. *Journal of Clinical Epidemiology*; 52(6): 569–83.

Rowan S, Rachet B, Alexe DM, Cooper N and Coleman MP (2008). 'Survival from prostate cancer in England and Wales up to 2001'. *British Journal of Cancer*; 99(Suppl.1): S75–S77.

Ruddon RW (2007). *Cancer Biology.* (Fourth edn., New York: Oxford University Press).

Rutherford MJ, Dickman PW and Lambert PC (in press). 'Comparison of methods for calculating relative survival in population-based studies'. *International Journal of Cancer Epidemiology, Detection and Prevention*; doi:10.1016/j.canep.2011.05.010.

Short PF, Vasey JJ and BeLue R (2008). 'Work disability associated with cancer survivorship and other chronic conditions'. *Psycho-Oncology*; 17(1): 91–97.

Simonelli C, Annunziata MA, Chimienti E, Berretta M and Tirelli U (2008). 'Cancer survivorship: a challenge for the European oncologists'. *Annals of Oncology*; 19(7): 1216–17.

Smastuen M, Aagnes B, Johannesen TB, Møller B and Bray F (2008). *Long-term cancer survival: patterns and trends in Norway 1965–2007.* (Oslo: Cancer Registry of Norway).

Smith MY, Redd WH, Peyser C and Vogl D (1999). 'Post-traumatic stress disorder in cancer: a review'. *Psycho-Oncology*; 8(6): 521–37.

Stenbeck M, Rosen M and Sparen P (1999). 'Causes of increasing cancer prevalence in Sweden'. *Lancet*; 354(9184): 1093–94.

Stigler SM (1999). 'Stigler's Law of Eponymy' *Statistics on the Table* (Cambridge, Massachusetts: Harvard University Press), 277–90.

Tabata N, Ohno Y, Matsui R, Sugiyama H, Ito Y, Tsukuma H and Oshima A (2008). 'Partial cancer prevalence in Japan up to 2020: Estimates based on incidence and survival data from population-based cancer registries'. *Japanese Journal of Clinical Oncology*; 38(2): 146–57.

Thames Cancer Registry (2011). 'Fifty years of cancer registration in South East England'. [Accessed 22 November, 2011]; http://www.thames-cancer-reg.org.uk/informat/pubs/50years_tcr_report.pdf.

Trent TT and Rongstad OJ (1974). 'Home Range and Survival of Cottontail Rabbits in Southwestern Wisconsin'. *The Journal of Wildlife Management*; 38(3): 459–72.

Twombly R (2004). 'What's in a name: who is a cancer survivor?'. *Journal of the National Cancer Institute*; 96(19): 1414–15.

US Department of Public Health (2011). 'Connecticut Tumour Registry'. [Accessed 22 November, 2011]; http://www.ct.gov/dph/cwp/view.asp?a=3129&q=389716.

van Leeuwen PJ, Connolly D, Tammela TL, Auvinen A, Kranse R, Roobol MJ, Schroder FH and Gavin A (2010). 'Balancing the harms and benefits of early detection of prostate cancer'. *Cancer*; 116(20): 4857–65.

Vandeschrick C (2001). 'The Lexis diagram, a misnomer'. *Demographic Research*; 4(3): 97–124.

Vera-Garcia C (2005). 'A patient's perspective on ovarian cancer as a chronic disease–Devaluing a survivor's challenge'. *Gynecologic Oncology*; 96(1): 19–20.

Verdecchia A, Capocaccia R, Egidi V and Golini A (1989). 'A method for the estimation of chronic disease morbidity and trends from mortality data'. *Statistics in Medicine*; 8(2): 201–16.

Verdecchia A, De Angelis G and Capocaccia R (2002). 'Estimation and projections of cancer prevalence from cancer registry data'. *Statistics in Medicine*; 21(22): 3511–26.

Verdecchia A, De Angelis R, Capocaccia R, Sant M, Micheli A, Gatta G and Berrino F (1998). 'The cure for colon cancer: results from the EUROCARE study'. *International Journal of Cancer*; 77(3): 322–29.

Verdecchia A, Mariotto A, Capocaccia R, Gatta G, Micheli A, Sant M and Berrino F (2001). 'Incidence and prevalence of all cancerous diseases in Italy: trends and implications'. *European Journal of Cancer*; 37(9): 1149–57.

Verdecchia A, Micheli A and Gatta G (eds.) (1997). 'Survival of cancer patients in Italy: the ITACARE study' in *Tumori;* 83: 1–507.

World Health Organisation (2005). 'Guidelines for measuring national HIV prevalence in population-based surveys'. [Accessed 22 November, 2011]; http://www.who.int/hiv/pub/surveillance/guidelinesmeasuringpopulation.pdf.

World Health Organisation (2010). 'ICD-10 Version 2010'. [Accessed 23 November, 2011]; http://apps.who.int/classifications/icd10/browse/2010/en.

Yabroff KR, Lamont EB, Mariotto A, Warren JL, Topor M, Meekins A and Brown ML (2008a). 'Cost of care for elderly cancer patients in the United States'. *Journal of the National Cancer Institute*; 100(9): 630–41.

Yabroff KR, Lawrence WF, Clauser S, Davis WW and Brown ML (2004). 'Burden of illness in cancer survivors: findings from a population-based national sample'. *Journal of the National Cancer Institute*; 96(17): 1322–30.

Yabroff KR, Mariotto AB, Feuer E and Brown ML (2008b). 'Projections of the costs associated with colorectal cancer care in the United States, 2000-2020'. *Health Economics*; 17(8): 947-59.

Zink A, Rohrbach H, Szeimies U, Hagedorn HG, Haas CJ, Weyss C, Bachmeier B and Nerlich AG (1999). 'Malignant tumors in an ancient Egyptian population'. *Anticancer Research*; 19(5B): 4273–77.

Appendix A. List of publications

The following journal articles featuring work contained in, or related to, this thesis were published between 2009 and 2011. Reprints of the articles for which I was the first author are contained in Appendix C.

Maddams J, Brewster D, Gavin A, Steward J, Elliott J, Utley M and Møller H (2009). 'Cancer prevalence in the United Kingdom: estimates for 2008'. *British Journal of Cancer*; 101(3), 541–47.

Maddams J, Utley M and Møller H (2011). 'Levels of acute health service use among cancer survivors in the United Kingdom'. *European Journal of Cancer*; 47(14), 2211–20.

Maddams J, Utley M and Møller H (2011). 'A person-time analysis of hospital activity among cancer survivors in England'. *British Journal of Cancer*; 105(Suppl.1), S38–45.

Fiorentino F, Maddams J, Møller H and Utley M (2011). 'Modelling to estimate future trends in cancer prevalence'. *Health Care Management Science*; 14(3), 262–66.

Appendix B. List of presentations 2008–2011

Maddams J, Utley M and Møller H. 'Cancer prevalence and survivorship in the Thames Cancer Registry region'. *Thames Cancer Registry, King's College London.* Oral presentation; London, England: May 2008.

Maddams J, Utley M and Møller H. 'Cancer survivors and cancer survivorship. Quantifying cancer prevalence and modelling its dynamics in England'. *UK Association of Cancer Registries Annual Conference. Keble College, University of Oxford.* Oral presentation; Oxford, England: September 2008.

Maddams J, Utley M and Møller H. 'Two million cancer survivors in the UK'. *National Cancer Research Institute Annual Conference.* Poster presentation; Birmingham, England: October 2008.

Maddams J, Utley M and Møller H. 'Living with and beyond cancer: Phases of survivorship and dynamics of cancer prevalence'. *Clinical Operational Research Unit, University College London.* Oral presentation; London, England: February 2009.

Maddams J, Utley M and Møller H. 'Cancer prevalence in the UK – estimates for 2008'. *International Association of Cancer Registries Annual Conference.* Poster presentation; New Orleans, US: May 2009.

Maddams J, Utley M and Møller H. 'Cancer prevalence in the UK – estimates for 2008'. *Division of Cancer Studies Research Showcase, King's College London.* Poster presentation; London, England: June 2009.

Maddams J, Utley M and Møller H. 'Cancer prevalence in the UK – estimates for 2008'. *UK Association of Cancer Registries Annual Conference.* Poster presentation; Falkirk, Scotland: September 2009.

Maddams J, Utley M and Møller H. 'A person-time approach to Hospital Episode Statistics data'. *Thames Cancer Registry, King's College London.* Oral presentation; London, England: March 2010.

Maddams J, Utley M and Møller H. 'Person-time analysis of hospital activity among cancer survivors'. *UK Association of Cancer Registries and National Cancer Intelligence Network Annual Conference.* Poster presentation; Birmingham, England: June 2010.

Maddams J, Utley M and Møller H. 'Person-time analysis of hospital activity among cancer survivors'. *Division of Cancer Studies Research Showcase, King's College London*. Poster presentation; London, England: June 2010.

Maddams J, Utley M and Møller H. 'Cancer registry data linked to hospital episode statistics'. *National Cancer Research Institute Annual Conference.* Oral presentation; Liverpool, England: November 2010.

Maddams J, Utley M and Møller H. 'Cancer survivors and cancer survivorship – quantifying cancer prevalence and modelling its dynamics in England and the UK'. *Macmillan Cancer Support*. Oral presentation; London, England: January 2011.

Maddams J and Møller H. 'Diabetes, co-morbidity and colorectal cancer survival'. *International Diabetes and Cancer Research Consortium*. Oral presentation; Banff, Canada: March 2011.

Maddams J, Utley M and Møller H. 'Cancer survivorship and acute health service utilisation in the UK'. *Research Oncology, King's College London.* Oral presentation; London, England: April 2011.

Maddams J, Utley M and Møller H. 'Acute health care utilisation among cancer survivors in the UK'. *UK Association of Cancer Registries and National Cancer Intelligence Network Annual Conference.* Oral presentation; London, England: June 2011.

Maddams J and Møller H. 'Diabetes, co-morbidity and cancer survival'. *UK Association of Cancer Registries and National Cancer Intelligence Network Annual Conference*. Poster presentation; London, England: June 2011.

Maddams J, Darby S, Parkin DM. 'The cancer burden in the UK in 2008 due to radiotherapy'. *UK Association of Cancer Registries and National Cancer Intelligence Network Annual Conference*. Poster presentation; London, England: June 2011.

Maddams J, Utley M and Møller H. 'Acute health care utilisation among cancer survivors in the UK'. *Division of Cancer Studies Research Showcase, King's College London*. Oral presentation; London, England: June 2011.

Maddams J, Utley M and Møller H. 'Cancer survivorship and acute health service utilisation in the UK'. *MSc in Palliative Care: Service Development and Management Module, King's College London*. Oral presentation; London, England: October 2011.

Appendix C. First author publication reprints

British Journal of Cancer (2009) 101, 541 – 547

www.bjcancer.com

Cancer prevalence in the United Kingdom: estimates for 2008

J Maddams*[,1], D Brewster[2], A Gavin[3], J Steward[4], J Elliott[5], M Utley[6] and H Møller[1]

[1]Kings College London, Thames Cancer Registry, 1st Floor Capital House, 42 Weston St, London SE1 3QD, UK; [2]Scottish Cancer Registry, Area 155, Gyle Square, 1 South Gyle Crescent, Edinburgh EH12 9EB, UK; [3]Northern Ireland Cancer Registry, Queen's University Belfast, School of Medicine Dentistry and Biomedical Sciences, Mulhouse Building, Grosvenor Road, Belfast BT12 6BJ, UK; [4]Welsh Cancer Intelligence and Surveillance Unit, Floor 13, Brunel House, 2 Fitzalan Road, Cardiff CF24 0HA; [5]Macmillan Cancer Support, 89 Albert Embankment, London SE1 7UQ, UK; [6]University College London, Clinical Operational Research Unit, 4 Taviton Street, London WC1H 0BT, UK

BACKGROUND: Identifying and addressing the requirements of cancer survivors is currently a high priority for the NHS, yet little is known about the population of cancer survivors in the United Kingdom.

METHODS: Data from cancer registries in England, Northern Ireland, Scotland and Wales were analysed to provide limited-duration prevalence estimates for 2004. Log-linear regression models were used to extend these to complete prevalence estimates. Trends in prevalence from 2000 to 2004 were used to project complete prevalence estimates forward from 2004 to 2008.

RESULTS: We estimated that in total, there were 2 million cancer survivors in the United Kingdom at the end of 2008, ~3% of the population overall and 1 in 8 of those aged 65 years and more. Prostate and female breast cancers were the most prevalent. The number of cancer survivors is increasing by ~3% each year. Estimates are also provided by time since diagnosis.

CONCLUSION: These estimates are the most up-to-date available, and as such will be useful for statutory and voluntary sector organisations that are responsible for planning and providing treatment and support to cancer survivors in the United Kingdom.

British Journal of Cancer (2009) **101,** 541 – 547. doi:10.1038/sj.bjc.6605148 www.bjcancer.com
Published online 30 June 2009

Keywords: prevalence; survivors; survivorship; UK

Cancer survivors are defined as those people living with a diagnosis of cancer from some point in their past. Cancer prevalence is expressed as the number or proportion of cancer survivors in a population at a given point in time. Identifying and addressing the requirements of cancer survivors in England is a high priority in the Cancer Reform Strategy (Department of Health, 2007) and, as a result, the National Cancer Survivorship Initiative was set up in 2008. Similar initiatives are being established in Northern Ireland, Scotland and Wales. However, little is known about the size and demography of the population of cancer survivors in the United Kingdom; the most recent estimate, provided by the EUROPREVAL project (Forman *et al*, 2003), was for 1992.

In this paper, we provide up-to-date estimates of cancer prevalence in the United Kingdom at the end of 2008 using cancer registry data. Cancer survivors may have been recently diagnosed and in active treatment, or they may have survived long enough to be considered cured. However, in our analysis, we do not make such distinctions; once an individual is diagnosed with cancer, he (or she) is considered a survivor until death. We adopt this approach because, first, a diagnosis of cancer may affect a person's life in different ways (mental health, fear of recurrence, financial hardship, relationship issues, etc.), and its effects may be felt for many years after diagnosis. Second, this approach is practical as the currently available cancer registration data do not readily allow survivors to be classified as having active disease, in remission or cured of their cancer.

*Correspondence: J Maddams, Thames Cancer Registry, 1st Floor, Capital House, 42 Weston St, London SE1 3QD, UK;
E-mail: jacob.maddams@kcl.ac.uk
Revised 20 May 2009; accepted 29 May 2009; published online 30 June 2009

MATERIALS AND METHODS

The eight cancer registries in England, together with the national registries in Scotland, Wales and Northern Ireland, provided anonymised records of all registered malignant neoplasms (ICD-10 C00-C97) diagnosed in the residents of those countries, excluding non-melanoma skin cancer (ICD-10 C44) as it is not covered systematically by all registries. Each record included demographic, tumour, diagnosis, follow-up and death details. Data were available for the periods 1971 – 2004 for England, 1971 – 2005 for Scotland, 1990 – 2006 for Wales and 1993 – 2006 for Northern Ireland. All tumours apparently diagnosed in patients over the age of 99 years were excluded (~0.04% of the total), leaving 7.7 million registration records for analysis.

The UK cancer registries receive death notifications from the Office for National Statistics (ONS) (England and Wales) and the General Register Offices (Scotland and Northern Ireland), which are then matched to the cancer registration records, although a small percentage are never so matched. The patients associated with these 'lost-to-follow-up' registrations are at face value, effectively immortal, resulting in apparent cancer survivors of a much higher age than we know to be likely. The proportion of registrations lost to follow-up in European registries is believed to be <1% (Capocaccia *et al*, 1999), but is unknown in the United Kingdom. Therefore, in computing prevalence, cancer survivors were censored at the age of 105 years.

Cancer prevalence can be expressed as the number of prevalent tumours or the number of prevalent patients. As each patient may,

in their lifetime, be diagnosed with more than one tumour, patient prevalence will always be lower than tumour prevalence. This analysis focuses on patient prevalence, and only the first diagnosed malignant neoplasm (excluding non-melanoma skin cancer) in each patient was considered.

Although cancer registry data for Scotland, Wales and Northern Ireland were available for the years 2005 and 2006, for England it was available only up to 2004 at the time of analysis. For this reason, we used the most recent index date common to all data, that is, 31 December 2004. The number of cancer patients alive on this date was counted and disaggregated by country of residence, sex, age group on the index date (0–44, 45–64 and 65+ years), number of years since diagnosis and the following broad groups of cancer diagnoses:

(1) colon, rectum and anus (ICD-10 C18-C21),
(2) lung, bronchus and trachea (ICD-10 C33-C34),
(3) female breast (ICD-10 C50),
(4) prostate (ICD-10 C61),
(5) all other malignant neoplasms excluding non-melanoma skin cancer (ICD-10 C00-C97 excluding C44 and (1) to (4)).

Death-Certificate-Only (DCO) registrations are those for which the only source of patient/tumour information is the death certificate stating the cause of death. These registrations lack much information, particularly the actual date of diagnosis. An unknown proportion of the DCO registrations since the index date will pertain to patients diagnosed before, and alive on, the index date. We have not attempted to estimate this proportion, and therefore they are not included in our prevalence estimates.

Complete prevalence

N-year limited-duration prevalence counts include only those survivors diagnosed in the last *N* years before the index date. Complete prevalence includes all cancer survivors, regardless of when they were diagnosed. It is not possible to directly count complete prevalence on the basis of registry data, given that no UK cancer registry has been collecting data for a sufficiently long period of time. Instead, we estimated complete prevalence by extrapolating from limited-duration prevalence.

With an index date of 31 December 2004, the available cancer registry data provided 34-year prevalence estimates for England and Scotland, 15-year estimates for Wales and 12-year estimates for Northern Ireland. To extend these limited-duration estimates to complete estimates, a negative binomial regression model with a log-link function was constructed for each type of cancer, sex and age group (0–44, 45–64, 65+ years). The prevalence count on the index date was the response variable, and the predictor variables were country of residence and number of years since diagnosis. Given that our primary objective was to obtain a reasonable estimate for the number of people surviving at least 12 years (Northern Ireland) and 15 years (Wales) beyond diagnosis, data pertaining to recent diagnoses (years since diagnosis $\leqslant 5$) were not used in the models. Prostate cancer was treated as a special case and modelled in two stages; first, for years of diagnosis between 1992 and 1999 and second, for all years before 1992. We also included an offset term in all models, defined as the log of the number of people in a given country who could contribute to the prevalence count, taking into account the age group being considered and the fact that years since diagnosis cannot exceed age on the index date. The models were run using the PROC GENMOD procedure in SAS (SAS Institute Inc, Cary, NC, USA).

The validity of the regression models was tested by initially excluding data for Scotland covering years of diagnosis between 1971 and 1992, and by comparing the modelled estimates with the actual data for those years. Furthermore, we compared the published estimated ratios of 15-year prevalence to complete prevalence (the completeness index) in England and Scotland (Forman *et al*, 2003) with those of our own (Table 1).

Table 1 Comparison of 15-year completeness indices, by cancer site and sex

	Maddams *et al*[a]	EUROPREVAL[b]
Males		
Colon, rectum and anus	0.81	0.87[c]
Lung, bronchus and trachea	0.58	0.73[d]
Prostate	0.95	0.97
All malignant neoplasms[e]	0.78	0.82
Females		
Breast	0.74	0.80
Colon, rectum and anus	0.74	0.80
Lung, bronchus and trachea	0.77	0.79
All malignant neoplasms[e]	0.70	0.72

ICD = International Classification of Diseases. The 15-year completeness index is defined as 15-year prevalence divided by complete prevalence. [a]15-year prevalence divided by estimated complete prevalence in the United Kingdom. Index date: 31 December 2004. [b]Average 15-year completeness index for prevalence as estimated using data from South Thames, West Midlands, Yorkshire and Scotland cancer registries. Index date: 31 December 1992. Published by Forman *et al* (2003), *Ann Oncol* 14: 648–654. [c]Average of indices for cancers of the colon and rectum. [d]Cancer of the lung only. [e]Excluding non-melanoma skin cancer (ICD-10 C44).

Trends

Through an analysis of recent trends in observed cancer prevalence, we projected estimates of complete cancer prevalence in the United Kingdom from 31 December 2004 forward to 2008. We used the combined data from England and Scotland covering diagnoses between 1971 and 2004 to estimate trends in limited-duration prevalence during the years 2000–2004, for each site and sex. Log-linear functions, considered appropriate for short-term projections, were fitted to the trend data and provided estimates of the annual growth in cancer survivor numbers that we expected from 2004 to 2008. The following assumptions were made:

(1) the yearly rates of change of cancer prevalence in England and Scotland combined can be applied to each country in the United Kingdom;
(2) cancer prevalence in each age group (0–44, 45–64 and 65+ years) is changing at the same rate as overall prevalence;
(3) complete prevalence is changing at the same rate as 30-year prevalence.

Estimated prevalence counts were converted to proportions of the population by using the mid-year ONS population estimates for 2007; these were the most recently available estimates and likely to be only slightly lower than the actual population at the end of 2008 (ONS, 2008a).

RESULTS

Tables 2 and 3 and Figure 2 present complete prevalence – the sum of observed prevalence from the years of diagnosis that were available in our data and modelled prevalence from those that were not. We have indicated by italicised text in Tables 2–4 those estimates for which more than 20% of the total is derived from modelling. We estimated that by the end of 2008, there were just over 2 million cancer survivors in the United Kingdom (59% women and 41% men), equating to $\sim$2.7% of the male and 3.8% of the female population. Table 2 shows the variation in prevalence by country and cancer sites. Wales had the most number of cancer

Epidemiology

Table 2 Prevalence of cancer on 31 December 2008 in the United Kingdom, by country of residence

	England	Scotland	Wales	Northern Ireland	United Kingdom
Males					
Colon, rectum and anus	100 608 (401)	11 522 (464)	6921 (476)	*3480 (404)*	122 531 (410)
Lung, bronchus and trachea	32 034 (128)	3760 (151)	*1889 (130)*	*1058 (123)*	*38 741 (130)*
Prostate	215 654 (859)	19 163 (771)	13 312 (916)	5307 (616)	253 436 (847)
All other malignant neoplasms[a]	334 147 (1330)	36 853 (1483)	*22 998 (1582)*	*10 482 (1216)*	404 480 (1352)
All malignant neoplasms[a]	682 443 (2717)	71 298 (2868)	*45 120 (3103)*	*20 327 (2358)*	819 188 (2738)
Females					
Breast	460 041 (1771)	46 211 (1738)	*29 838 (1955)*	*12 908 (1439)*	548 998 (1768)
Colon, rectum and anus	92 439 (356)	11 419 (430)	*5885 (386)*	*3542 (395)*	113 285 (365)
Lung, bronchus and trachea	19 634 (76)	3215 (121)	1239 (81)	*693 (77)*	24 781 (80)
All other malignant neoplasms[a]	409 284 (1576)	46 607 (1753)	*26 683 (1749)*	*13 690 (1526)*	496 264 (1598)
All malignant neoplasms[a]	981 398 (3778)	107 452 (4042)	*63 645 (4171)*	*30 833 (3437)*	1 183 328 (3810)

ICD = International Classification of Diseases. The number of survivors (crude proportion per 100 000) is indicated; the sum of observed prevalence was available from cancer registry data, whereas modelled prevalence was not. Italicised numbers are those that are based on estimates of prevalence in 2004 that were at least 20% modelled. [a]Excluding non-melanoma skin cancer (ICD-10 C44).

Table 3 Prevalence of cancer on 31 December 2008 in the United Kingdom, by time since diagnosis

	0–1 years	1–5 years	5–10 years	10–20 years	>20 years	Total
Males						
Colon, rectum and anus	14 619 (49)	38 075 (127)	31 162 (104)	24 534 (82)	*14 141 (47)*	122 531 (410)
Lung, bronchus and trachea	8263 (28)	7850 (26)	3810 (13)	4769 (16)	*14 049*[a] *(47)*[a]	*38 741 (130)*
Prostate	37 967 (127)	125 470 (419)	61 376 (205)	22 601 (76)	6022 (20)	253 436 (847)
All other malignant neoplasms[b]	40 891 (137)	97 529 (326)	87 008 (291)	98 726 (330)	*80 326 (269)*	404 480 (1352)
All malignant neoplasms[b]	101 740 (340)	268 924 (899)	183 356 (613)	150 630 (504)	*114 538 (383)*	819 188 (2738)
Females						
Breast	42 432 (137)	140 111 (451)	128 672 (414)	145 035 (467)	*92 748 (299)*	548 998 (1,768)
Colon, rectum and anus	11 309 (36)	30 341 (98)	27 128 (87)	25 532 (82)	*18 975 (61)*	113 285 (365)
Lung, bronchus and trachea	6905 (22)	7255 (23)	3671 (12)	2682 (9)	*4268 (14)*	24 781 (80)
All other malignant neoplasms[b]	40 655 (131)	109 179 (352)	96 400 (310)	119 366 (384)	*130 664 (421)*	496 264 (1598)
All malignant neoplasms[b]	101 301 (326)	286 886 (924)	255 871 (824)	292 615 (942)	*246 655 (794)*	1 183 328 (3810)

ICD = International Classification of Diseases. The number of survivors (crude proportion per 100 000) is indicated; the sum of observed prevalence was available from cancer registry data, whereas modelled prevalence was not. Italicised numbers are those that are based on estimates of prevalence in 2004 that were at least 20% modelled. [a]The estimate of the number of long-term male lung cancer survivors (>20 years from diagnosis) is likely over-estimated; see the section 'Discussion' for details. A more plausible figure is 6000 (20 per 100 000). [b]Excluding non-melanoma skin cancer (ICD-10 C44).

Table 4 Prevalence of cancer on 31 December 2008 in the United Kingdom, by age

	0–44 years	45–64 years	65+ years	Total
Males				
Colon, rectum and anus	2091 (11)	25 690 (343)	94 750 (2238)	122 531 (410)
Lung, bronchus and trachea	441 (2)	6643 (89)	*31 657 (748)*	*38 741 (130)*
Prostate	181 (1)	34 511 (461)	218 744 (5168)	253 436 (847)
All other malignant neoplasms[a]	68 539 (377)	125 077 (1671)	210 864 (4982)	404 480 (1352)
All malignant neoplasms[a]	71 252 (392)	191 921 (2563)	556 015 (13 136)	819 188 (2738)
Females				
Breast	25 428 (143)	208 076 (2694)	315 494 (5688)	548 998 (1768)
Colon, rectum and anus	2134 (12)	19 723 (255)	91 428 (1648)	113 285 (365)
Lung, bronchus and trachea	530 (3)	5904 (76)	18 347 (331)	24 781 (80)
All other malignant neoplasms[a]	67 530 (380)	151 756 (1965)	276 978 (4994)	496 264 (1598)
All malignant neoplasms[a]	95 622 (538)	385 459 (4990)	702 247 (12 661)	1 183 328 (3810)

ICD = International Classification of Diseases. The number of survivors (crude proportion per 100 000) is indicated; the sum of observed prevalence was available from cancer registry data, whereas modelled prevalence was not. Italicised numbers are those that are based on estimates of prevalence in 2004 that were at least 20% modelled. [a]Excluding non-melanoma skin cancer (ICD-10 C44).

survivors per capita (3.1% men and 4.2% women), and Northern Ireland had the fewest (2.4% men and 3.4% women).

Prostate and female breast cancers were the most prevalent, and accounted for 31 and 46% of male and female cancer prevalence, respectively. Of the cancers studied in this paper, lung cancer was the least prevalent. Figure 1 shows, for each sex, the proportions of total incident cases, cancer deaths and cancer survivors that are accounted for by colorectal, lung, prostate and female breast

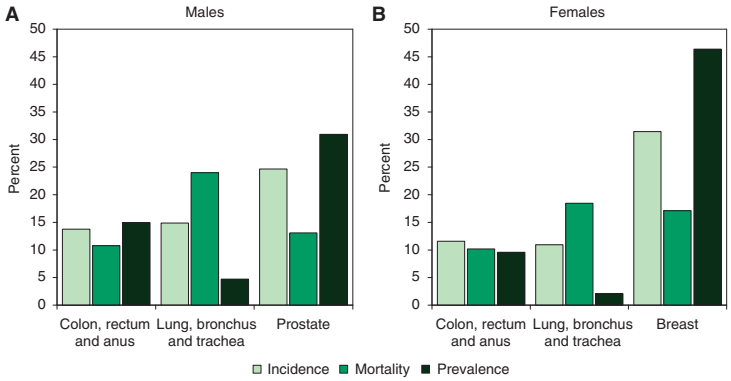


Figure 1 Proportion of total incidence[1], mortality[2] and prevalence[3] that is accounted for by selected cancers. (**A**) Males; (**B**) females. [1]Incidence in England, 2006; data from National Cancer Information Service (NCIS); [2]Mortality in England, 2005; data from NCIS; [3]Prevalence in the United Kingdom, 2008.

cancers. For both men and women, colorectal cancer accounted for approximately 10–15% of all the three measures. In contrast, for men, lung cancer accounted for 15% of all newly diagnosed cancers, 25% of cancer deaths and for only 5% of cancer prevalence. A similar pattern was seen for female lung cancer, which accounted only for 2% of cancer prevalence in women. Prostate and female breast cancers provided further contrasts, the latter accounting for 31% of newly diagnosed cancers, 17% of cancer deaths and for 46% of cancer prevalence among women.

Table 3 presents cancer prevalence in the United Kingdom, by the number of years that had passed since diagnosis. This varied across the sexes and cancer sites, as illustrated in Figure 2. Overall, female cancer survivors tended to be further from their diagnosis than males, 67% of them being diagnosed more than 5 years earlier, compared with 55% of males.

Table 4 shows the variation of cancer prevalence with age. Less than 1% of the UK population aged <45 years at the end of 2008 were cancer survivors, compared with 13% of those aged 65 years and more. There were twice as many female survivors aged between 45 and 64 years as there were males, largely because of the dominance of female breast cancer that accounted for 54% of female survivors in this age range. The most prevalent types of cancer in those aged 65 years and more were prostate and female breast cancers; 5% of males and 6% of females in this age group were survivors of these cancers.

Figure 3 shows the trends in 1-, 5-, 10-, 20- and 30-year limited-duration prevalence during the period 2000–2004, projected to 2008. Only male lung cancer did not show an increasing trend, the total number of survivors declining by 1.4% per year. By far, the most rapidly increasing prevalence is that of prostate cancer, the total number of survivors increasing by 9.8% per year. Overall, the number of cancer survivors increased by 3.8% per year for men and 2.7% for women.

Epidemiology

DISCUSSION

For producing our prevalence estimates, we have, where available, used incidence and follow-up data collected by cancer registries in the United Kingdom. We have not adjusted our estimates for DCO registrations that occurred after the index date. DCOs account for <5% of all registrations in the UK and most often relate to patients who have died soon after diagnosis. We have therefore assumed that their effect on cancer prevalence is negligible. We have not attempted to estimate the proportion lost to follow-up nor emigrations. Nor have we been able to include UK immigrants with a diagnosis of cancer pre-dating their move. To a certain extent, the effects of including emigrants and excluding immigrants will cancel each other out.

For estimating the number of survivors from a period before cancer registration in their country, we have developed log-linear regression models for the prevalence count as a function of time since diagnosis. Treating prevalence in this manner as an isolated statistic does not explicitly model the joint effect of incidence and survival, and is based on the observation that the relationship between the number of years since diagnosis and the number of prevalent cases is approximately log linear (Phillips *et al*, 2002). However, as this relationship is not log linear for prostate cancer, the regression model was applied in two stages. This accounted for the introduction of PSA testing in the early 1990s, which effectively changed the definition of the disease with many more localised tumours diagnosed (Evans and Moller, 2003). The number of prostate cancer survivors is increasing at the fastest rate of cancers studied here, by almost 10% each year. As the changes in incidence and survival caused by the introduction of PSA testing are relatively recent, we expect their numbers to increase at a similar rate for some years to come, until a situation is reached in which very few were diagnosed in the era before PSA testing.

Table 1 shows that our estimated 15-year completeness indices for 2004 were consistently lower than those previously published for 1992 (Forman *et al*, 2003). The differences in the completeness index are consistent with the improvements in survival observed between 1992 and 2004 for many cancers (Cancer Research UK, 2007; Rachet *et al*, 2008). Increases in survival rates result in a larger proportion of long-term cancer survivors and therefore, in a smaller completeness index. Although most differences in these indices were small, for male lung cancer, the difference was large (0.58 compared with the published estimate of 0.73). Despite the two estimates relating to different index dates and the changes in lung cancer incidence between the two dates, we would not expect such a large discrepancy. For lung cancer, as with prostate cancer, the number of survivors is not a log-linear function of years since diagnosis. Owing to its poor prognosis, lung cancer prevalence is dominated by short-term survivors (Figure 2c); hence, we believe that our models have over-estimated the number of long-term male survivors. If we were to use a ratio of 0.73, then we would produce an estimate of 6000 (rather than 14 000) male lung cancer

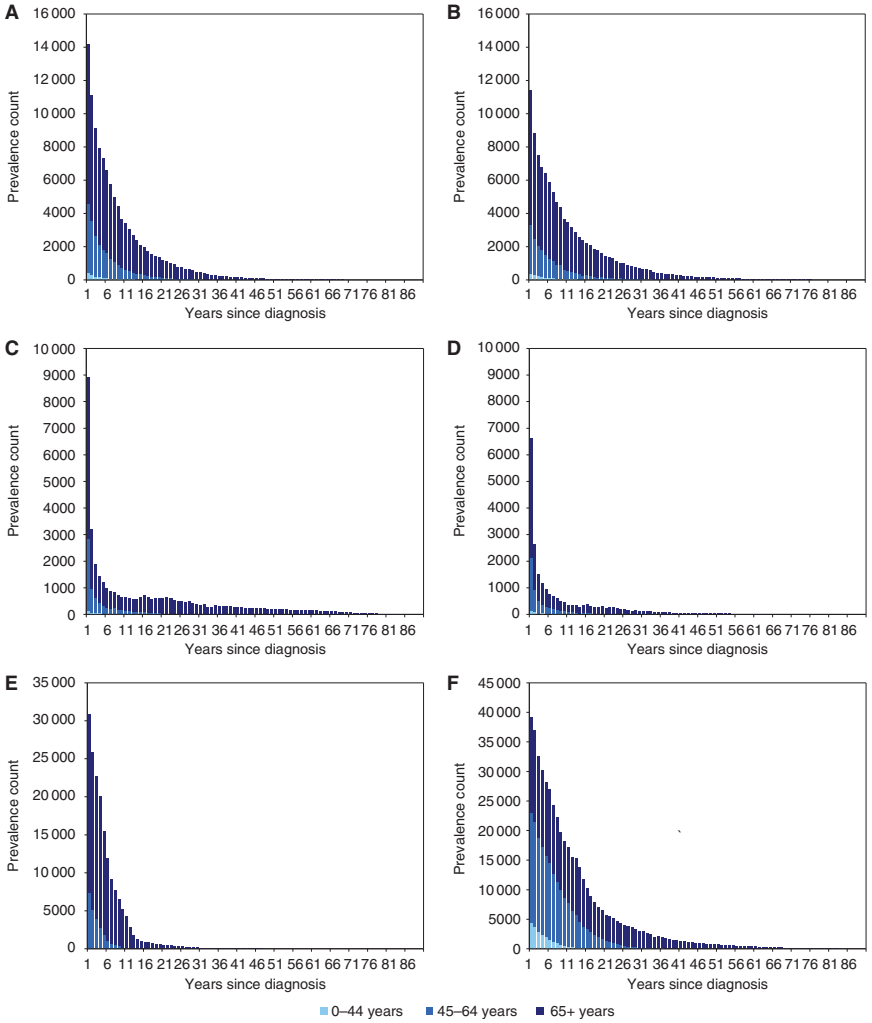


Figure 2 Prevalence of cancer in the United Kingdom on 31 December 2004, by number of years since diagnosis and age. (**A**) Males: colon, rectum and anus; (**B**) females: colon, rectum and anus; (**C**) males: lung, bronchus and trachea; (**D**) females: lung, bronchus and trachea; (**E**) prostate; (**F**) females: breast. Prevalence of cancer in the United Kingdom on 31 December 2004: the sum of observed prevalence was available from national cancer registry data, whereas modelled prevalence was not.

survivors who, at the end of 2008, have survived more than 20 years since their diagnosis. This estimate seems more plausible, especially when compared with the report stating that, in the Northern European countries in 1992, only 31% of male lung cancer prevalence was accounted for by survivors who were more than 10 years from diagnosis Moller *et al* (2003).

We have adopted a pragmatic approach to the modelling used in this study. Owing to the long time series of registry data in the United Kingdom, the majority of the estimates presented contain only a small contribution of modelled data. This contribution is most significant in the estimates for Northern Ireland and, to a lesser extent, for Wales. With 34-years of data available for England and Scotland, the modelled proportions of the complete prevalence estimates were typically ~5% for men and 8% for women. Therefore, we believe these estimates to be robust and fit for the purpose.

Substantial results of the analysis

There are ~2 million cancer survivors in the United Kingdom today. In those aged 65 years and more, ~13% of the population are cancer survivors. Approximately one in three of the UK

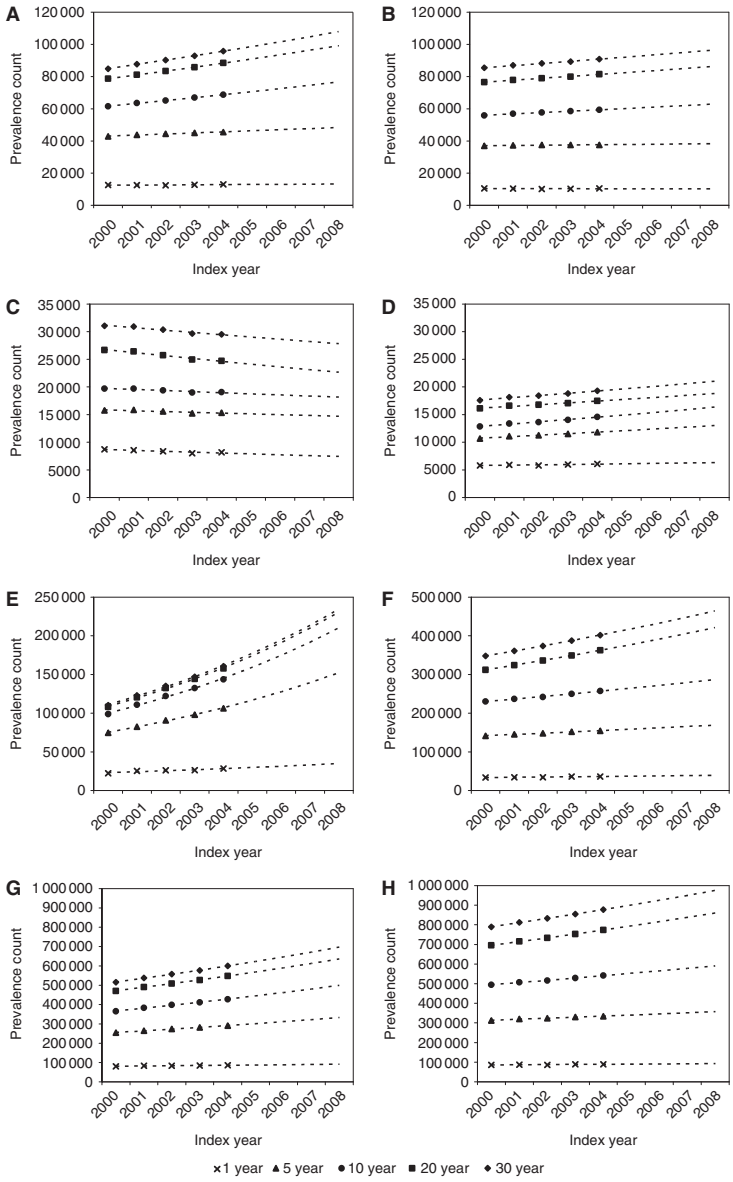


Figure 3 Trends and projections of limited-duration cancer prevalence in the United Kingdom, 2000–2008. (**A**) Males: colon, rectum and anus; (**B**) females: colon, rectum and anus; (**C**) males: lung, bronchus and trachea; (**D**) females: lung, bronchus and trachea; (**E**) prostate; (**F**) females: breast; (**G**) All malignant neoplasms[1]; males. (**H**) All malignant neoplasms[1]; females; [1]Excluding non-melanoma skin cancer (ICD-10 C44).

Epidemiology

population will be diagnosed with cancer during their lifetime and one in four will die from it (Cancer Research UK, 2008); we can also state that one in eight of those aged 65 years and more are living with or beyond cancer.

Our overall estimate of 2 million is far higher than that of 1.2 million at the end of 1992 (Forman *et al*, 2003). However, we have found that in recent years the absolute number of cancer survivors in the United Kingdom has increased by ~3% per annum, and that

if a similar rate of increase is assumed to apply over the entire period between 1992 and 2008, then the two figures are consistent. Not only is cancer prevalence increasing overall but the relative prevalence of different types of cancer is also changing. For example, prostate and female breast cancers have shown some of the largest increases in incidence and improvements in survival of all cancers in the United Kingdom since 1992 (Cancer Research UK, 2007), resulting in the proportion of total sex-specific prevalence accounted for by each increasing from 14 and 37% in 1992 (Forman *et al*, 2003) to 31 and 46% in 2008, respectively.

The number of cancer survivors varies within the United Kingdom, with Northern Ireland having the lowest prevalence proportion and Wales the highest. We have not presented age-specific proportions for each country, but the observed differences are, at least in part, attributable to the different age structures in each country; Northern Ireland has the youngest population (63% aged under 45 years; UK average of 59%) and Wales has the oldest population (18% aged 65 years and more; UK average of 16%). Different patterns of adopting PSA testing in the early 1990s resulted in higher detection rates of prostate cancer in Britain, compared with those in Northern Ireland. Consequently, recorded prostate cancer incidence between 1993 and 2003 remained lower in Northern Ireland than in the rest of the United Kingdom (Fitzpatrick *et al*, 2006).

In areas where cancer registration is less comprehensive than in the United Kingdom, models of incidence and survival rates have been developed to estimate cancer prevalence (Capocaccia and De Angelis, 1997). Owing to the large amount of cancer registry data available, little modelling was required, and for simplicity our approach treated prevalence as an isolated statistic. Nevertheless, it is important to appreciate that it is not an isolated measure, and that historical incidence and survival figures combine to produce the prevalence figures existing today. Figure 1 provides an illustration of this interaction: a cancer such as lung cancer, with a poor prognosis accounts for a very small proportion of prevalent cases, despite being one of the most commonly diagnosed cancers. Conversely, prostate and female breast cancers, with relatively good prognoses, account for larger proportions of cancer survivors than they do for new incident cases. Changes in incidence and survival (brought about by changes in lifestyle, population structure and health-service policy) will therefore have significant consequences for the prevalence of cancer. For example, since the 1970s, the number of smokers in the UK male population decreased from $\sim$50% to $\sim$20% (ONS, 2008b), resulting in a decrease in male lung cancer incidence and, in turn, a decrease in prevalence.

The previous most recent UK estimates of cancer prevalence were related to 1992 (Forman *et al*, 2003), since then, as we have shown, cancer prevalence has changed markedly. Therefore, our estimates are highly relevant for both statutory and voluntary sector organisations that are responsible for planning and providing treatment and support to cancer survivors in the United Kingdom. In the coming years, cancer prevalence will continue to increase as a result of the growing and ageing population of the United Kingdom, increased detection of cancer and improving survival rates. We estimated that, overall, the annual rate of increase in the number of cancer survivors is currently $\sim$3%, and we anticipate that this rate of increase will continue in the near future. Issues surrounding care and support for cancer survivors should, therefore, remain high on the public health agenda, with analysis and projections of cancer incidence, prevalence and mortality becoming increasingly central to resource-planning decisions. Cancer detection and treatment resources tend to focus on the most commonly diagnosed types of cancer, or major killers, but for cancer survivors, the most significant cancers are those with both a high incidence rate and a relatively good prognosis (such as prostate cancer and female breast cancer).

Knowledge of the natural progression of a particular type of cancer, together with the analysis of prevalence by time since diagnosis, gives some indication of different phases of survivorship, but does not fully show the extent to which survivors require, or are receiving, care and support. Many survivors will be newly diagnosed and in active treatment, others may be in a state of remission or recurrence with or without late effects of treatment, others may be receiving palliative care, whereas some may consider themselves to be completely free of cancer. Awareness of these differences is important when assessing the health-care burden of cancer, the financial costs of which are the highest during initial treatment and end of life care (Brown *et al*, 1999). We intend to extend the work presented in this paper in order to provide additional details about the different phases of survivorship and to analyse the factors that influence cancer prevalence over time.

ACKNOWLEDGEMENTS

The data for this analysis are from the cancer registries in the United Kingdom (http://www.ukacr.org, http://www.ncin.org.uk). Macmillan Cancer Support funded this research. The Northern Ireland Cancer Registry is funded by the Department of Health, Social Services and Public Safety Northern Ireland. The Welsh Cancer Intelligence and Surveillance Unit is funded by the Welsh Assembly Government under a service level agreement, and the First Minister owns the registry database.

REFERENCES

Brown ML, Riley GF, Potosky AL, Etzioni RD (1999) Obtaining long-term disease specific costs of care: application to Medicare enrollees diagnosed with colorectal cancer. *Med Care* **37:** 1249 – 1259

Cancer Research UK (2007) Trends in cancer survival rates. http://info.cancerresearchuk.org/cancerstats/survival/fiveyear/

Cancer Research UK (2008) CancerStats: Key Facts http://publications.cancerresearchuk.org/WebRoot/crukstoredb/CRUK_PDFs/AllCancersKeyFactsJune08.pdf (Accessed 17 June 2009)

Capocaccia R, De Angelis R (1997) Estimating the completeness of prevalence based on cancer registry data. *Stat Med* **16:** 425 – 440

Capocaccia R, Gatta G, Chessa E, Valente F (1999) The EUROCARE-2 study. *IARC Sci Publ* **151:** 1 – 40

Department of Health (2007) Cancer Reform Strategy http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_081006 (Accessed 17 June 2009)

Evans HS, Moller H (2003) Recent trends in prostate cancer incidence and mortality in southeast England. *Eur Urol* **43:** 337 – 341

Fitzpatrick D, Gavin A, Donnelly D (2006) *Cancer Trends in Northern Ireland: 1993 – 2003*. Northern Ireland Cancer Registry http://www.qub.ac.uk/research-centres/nicr/FileStore/PDF/Filetoupload,32872,en.pdf (Accessed 17 June 2009)

Forman D, Stockton D, Moller H, Quinn M, Babb P, De Angelis R, Micheli A (2003) Cancer prevalence in the UK: results from the EUROPREVAL study. *Ann Oncol* **14:** 648 – 654

Moller T, Anderson H, Aareleid T, Hakulinen T, Storm H, Tyggvadottir L, Corazziari I, Mugno E (2003) Cancer prevalence in Northern Europe: the EUROPREVAL study. *Ann Oncol* **14:** 946 – 957

Office for National Statistics (ONS) (2008a) Population Change: UK population increases by 388 000. http://www.statistics.gov.uk/cci/nugget.asp?ID = 950

Office for National Statistics (2008b) Smoking habits in Great Britain. http://www.statistics.gov.uk/cci/nugget.asp?id = 313

Phillips N, Coldman A, McBride ML (2002) Estimating cancer prevalence using mixture models for cancer survival. *Stat Med* **21:** 1257 – 1270

Rachet B, Woods LM, Mitry E, Riga M, Cooper N, Quinn MJ, Steward J, Brenner H, Esteve J, Sullivan R, Coleman MP (2008) Cancer survival in England and Wales at the end of the 20th century. *Br J Cancer* **99**(Suppl 1)**:** S2 – S10

EUROPEAN JOURNAL OF CANCER 47 (2011) 2211–2220

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejconline.com

Levels of acute health service use among cancer survivors in the United Kingdom

Jacob Maddams [a,*], *Martin Utley* [b], *Henrik Møller* [a]

[a] *King's College London, Thames Cancer Registry, London SE1 3QD, UK*
[b] *University College London, Clinical Operational Research Unit, UK*

ARTICLE INFO

Article history:
Available online 16 May 2011

Keywords:
Cancer survivorship
Cancer prevalence
Health service use
Cancer burden

ABSTRACT

Background: In the United Kingdom, there are approximately two million cancer survivors (3.2% of the entire population), composed of groups of people in different phases of survivorship and with different health service needs. The aim of this study was to quantify the level of acute health service utilisation by cancer survivors in the UK, according to tumour type, age, sex, time since diagnosis, and time until death.

Methods: Linked national cancer registry and hospital activity data were analysed. The data covered all cancer-related admissions to public hospitals operated by the National Health Service in England occurring in 2006 among people diagnosed with cancer in the period 1990–2006. The intensity of cancer-related health service utilisation was categorised as 'none', 'low' (up to 10% of an individual's time), or 'high' (>10% of an individual's time), among groups defined by time since diagnosis and time until death. Results were extrapolated from the population of England in 2006 (51 million) to that of the UK in 2008 (61 million).

Findings: Sixty one thousand of the two million cancer survivors (3%) were in the 'high' utilisation category; 240,000 (12%) were in the 'low' category; 1.70 million (85%) had no cancer-related hospital admissions. 147,000 cancer survivors (7%) were in the last year of their life, and it was this group that had the highest levels of hospital utilisation. 1.57 million cancer survivors (78%) were more than 1 year from both diagnosis and death, and had no cancer related hospital admissions.

Interpretation: A considerable proportion of cancer survivors in the UK have a high level of hospital utilisation soon after diagnosis or before death, but the large majority of them are neither recently diagnosed nor near the end of their life, and do not utilise acute health services for cancer-related care.

1. Introduction

Cancer prevalence, the proportion of a population that has been diagnosed with cancer, is one measure of the cancer burden on society. The prevalent population comprises 'cancer survivors' – those people alive following a diagnosis of cancer from some point in their past. The broad definition of cancer survivors includes those recently diagnosed, in active treatment, in remission, receiving treatment for recurrence, in end of life care, and those who are cured.

Previous work has shown that there are approximately two million cancer survivors in the United Kingdom and that in recent years this number has increased by approximately 3% per annum,[1] largely due to an ageing population, earlier

* *Corresponding author:* Address: King's College London, Thames Cancer Registry, 1st Floor Capital House, 42 Weston Street, London SE1 3QD, UK. Tel.: +44 207 378 7688x232; fax: +44 207 378 9510.
E-mail addresses: jacob.maddams@kcl.ac.uk (J. Maddams), m.utley@ucl.ac.uk (M. Utley), henrik.moller@kcl.ac.uk (H. Møller).
0959-8049/$ - see front matter
doi:10.1016/j.ejca.2011.04.015

detection of cancers, and improved treatment regimes leading to increased survival. The Cancer Reform Strategy[2] led to the establishment in 2008 of a National Cancer Survivorship Initiative to focus on the needs of the growing population of survivors in England. As observed cancer survival increases, so does the number of long term survivors. For many people diagnosed with cancer, it is no longer considered the death sentence it once was.[3,4] Indeed, it has been argued that in many cases cancer may be better described as a chronic illness, i.e. one characterised by a prolonged duration and a recurring nature.[5] It is useful therefore to go beyond a simple enumeration of survivors, since this tells us only a limited amount about the cancer burden in terms of (a) the personal psychosocial and physical burden to the individual survivor; and (b) the resource implications to the health service and society at large.

Prevalence is often estimated directly from cancer registry data, and it is easy to disaggregate the prevalence estimates according to year of diagnosis or equivalently time since diagnosis.[6–13] Intuitively this appears a good way to broadly categorise cancer survivors – certainly the first year following diagnosis is likely to be one of the most traumatic physically and emotionally. However, in the medium and long term after diagnosis, survivorship is likely to vary greatly across individuals, and it may not be sufficient to classify the population of cancer survivors simply by time since diagnosis.

This work presents an analysis of a linked cancer registry and health service activity dataset for England. A set of survivorship 'states' are defined in terms of four temporal 'phases of survivorship' (using time since diagnosis and time until death) and intensity of acute health service utilisation. Using person-time methods, the number of survivors in each of these states at the end of 2008 is estimated. Results are presented in a novel way using bespoke graphics.

2. Methods

2.1. Data

The analysis was based on two linked datasets. The English national merged cancer registry dataset, which featured patient and diagnostic information relating to all cancers diagnosed between 1990 and 2006 and recorded by the eight regional population-based cancer registries in England, was linked at the patient level to the English national Hospital Episode Statistics (HES) dataset[14] which contained patient, clinical, and administrative details for in-patients and day case patients treated in any hospital operated by the 166 National Health Service (NHS) providers in England. The linkage algorithm was 'rules-based' and used national health service number (which is unique to each patient), date of birth, date of death where appropriate, sex and postcode of residence. Each HES record defines a complete episode of care under a given consultant in a given NHS facility, and a patient's journey from admission to discharge may comprise many such episodes.

The national cancer registry dataset was used to define a cohort of cancer survivors who had been diagnosed with a malignant neoplasm (other than non-melanoma skin cancer) in the period 1990–2006, and were alive for at least some portion of 2006. Sub-cohorts were defined according to type of cancer: colon, rectum, and anus cancers (ICD-10 C18-C21); lung, bronchus, and trachea cancers (ICD-10 C33-C34); prostate cancer (ICD-10 C61); and female breast cancer (ICD-10 C50). Details of all hospital episodes of care occurring in 2006 for these cancer survivors were extracted from the linked dataset. Episodes were included in the analysis if they were 'cancer related', in that at least one of the 14 recorded hospital diagnosis codes was for a malignancy other than non-melanoma skin cancer.

Fig. 1 – Survivors in temporal phases A–D on 31st December 2006. dx = diagnosis; † = death; A–D = temporal phase of survivorship as defined in Box 1.

2.2. *Survivorship states*

Four distinct temporal phases of survivorship (Box 1 and Fig. 1) were defined according to time since diagnosis and time until death:

Box 1: temporal phases of survivorship	
Phase A	Less than 1 year from death and less than 1 year from diagnosis
Phase B	Less than 1 year from death and more than 1 year from diagnosis
Phase C	More than 1 year from death and less than 1 year from diagnosis
Phase D	More than 1 year from death and more than 1 year from diagnosis

For each survivor, the person-time for which they were prevalent in 2006 (i.e. that which was post-diagnosis and pre-death and overlapped the calendar year 2006) was split into segments according to the time points at which they moved between temporal phases or broad age bands (chosen to be 0–44, 45–64, and ⩾65 years). Since death notifications were available for survivors in the cancer registry dataset up to the end of 2008, it was possible to identify the points in time during 2006 at which any survivors entering the last year of their life did so. This person-time splitting procedure was executed using the SAS software package (SAS Institute Inc., Cary, NC, United States of America) and a series of programmes developed by JM for this task, similar to the programme *Lexis.sas* by Carstensen.[15]

Intensity of acute health service utilisation was defined in each segment of person-time according to the combined duration of cancer related hospital episodes that occurred. It was considered to be "high" if hospital activity accounted for more than 10% of the person-time segment; it was considered to be "low" if hospital activity accounted for some, but not more than 10%, of the segment; and a separate category was reserved for those segments of person-time which contained no hospital activity. There were, therefore, 12 possible survivorship states defined according to temporal phase and intensity of acute health service utilisation (Fig. 2). The total amount of person-time spent by the population of survivors in each of these states was calculated separately for each tumour group and sex.

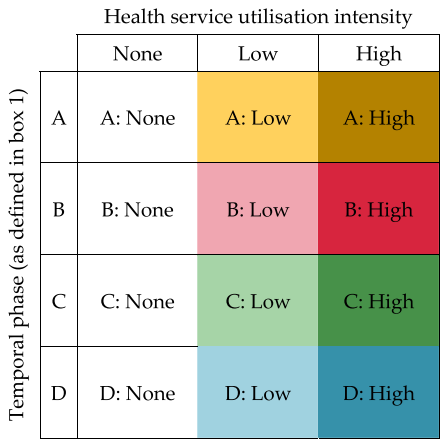


Fig. 2 – Survivorship states defined by time since diagnosis, time until death and health service utilisation. Key to Figs. 3 and 4.

Under certain assumptions, the proportion of total population person-time spent in each survivorship state is equal to the proportion of the prevalent population in that state at a given point in time during the analysis period (see Supplementary web extra material). Using this fact, the summed population person-time in our analysis was used to estimate the proportion of survivors in each state at the end of 2006.

2.3. *Extrapolation to complete prevalence in 2008*

In order to use our person-time analysis for the year 2006 to estimate the proportion of survivors in each survivorship state at the end of 2008 (the most recent estimate of prevalence in the UK[1]), it was assumed that the distribution of survivors between states did not change between 2006 and 2008.

The cancer registry dataset contained diagnoses made in the period 1990–2006. The person-time analysis for 2006 therefore completely described only those survivors diagnosed up to 16 years previously (16-year prevalence). By definition, survivors more than 16 years from diagnosis are in either temporal phase B or D. It was assumed that the distribution of survivors between these two phases was the same for those more than 16 years from diagnosis as it was for those between 15 and 16 years from diagnosis. It was also assumed that the relative numbers of survivors with an acute health service utilisation intensity of "high", "low", or "none" in each temporal phase B or D were the same for survivors greater than 16 years from diagnosis as they were for those in that phase no more than 16 years from diagnosis.

3. Results

Results are presented graphically in Figs. 3 and 4. Each figure contains a set of square tiles, with each tile representing cancer survivors of a given cancer type, sex, and broad age group. The density of the overlaid male or female icons is proportional to the total number of survivors represented by the tile. Four horizontal divisions mark the proportion of survivors in each of the four temporal phases (Box 1) and, in each phase, three vertical divisions mark the proportion of survivors in each category of acute health service utilisation. The resulting 12 areas are coloured to indicate the different survivorship states represented (Fig. 2).

These figures aim to present a large amount of data in a clean and simple manner, allowing for an immediate high level analysis, as well as facilitating in-depth study and easy comparison across cancer types, age groups, and sexes. The design was inspired by the work of the Dutch artist Piet Mondrian (1872–1944) who simplified visual compositions to three primary colours and a grid of black lines on a white background.[16] The underlying counts of survivors are given in Table 1, and the corresponding proportions in Table 2.

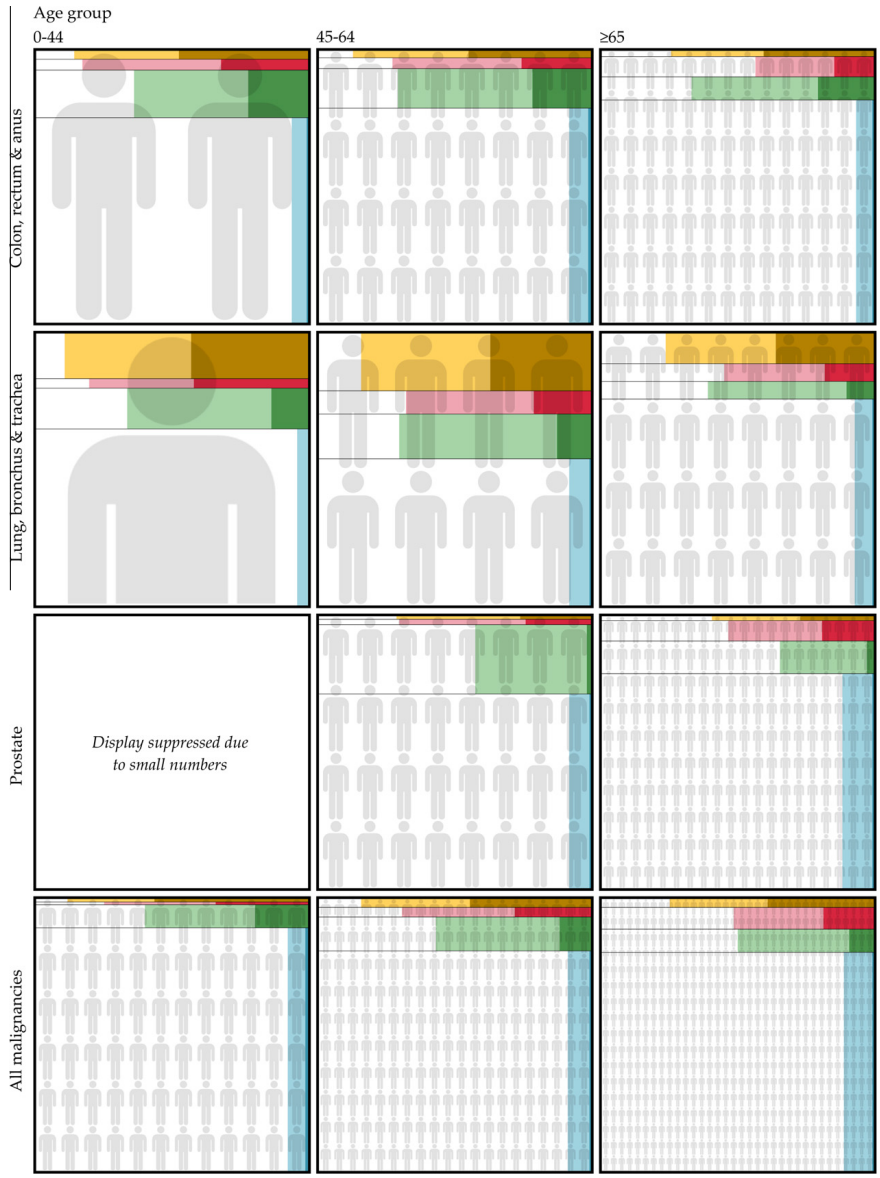


Fig. 3 – Male cancer survivors, UK, 2008; by age, type of cancer and survivorship state. In each tile, the rows from top to bottom correspond to the 4 temporal phases of survivorship (A-D) respectively, with the vertical bands from left to right corresponding to states of increasing acute health service utilisation (see Fig. 2 and Box 1 for definitions). Each male icon represents approximately 1,000 cancer survivors (exact counts are in Table 1). The proportion of survivors in temporal phase D (more than 1 year from both diagnosis and death) and with a 'high' level of acute health service utilisation (the rightmost area of the bottom row) may be too small to be visible.

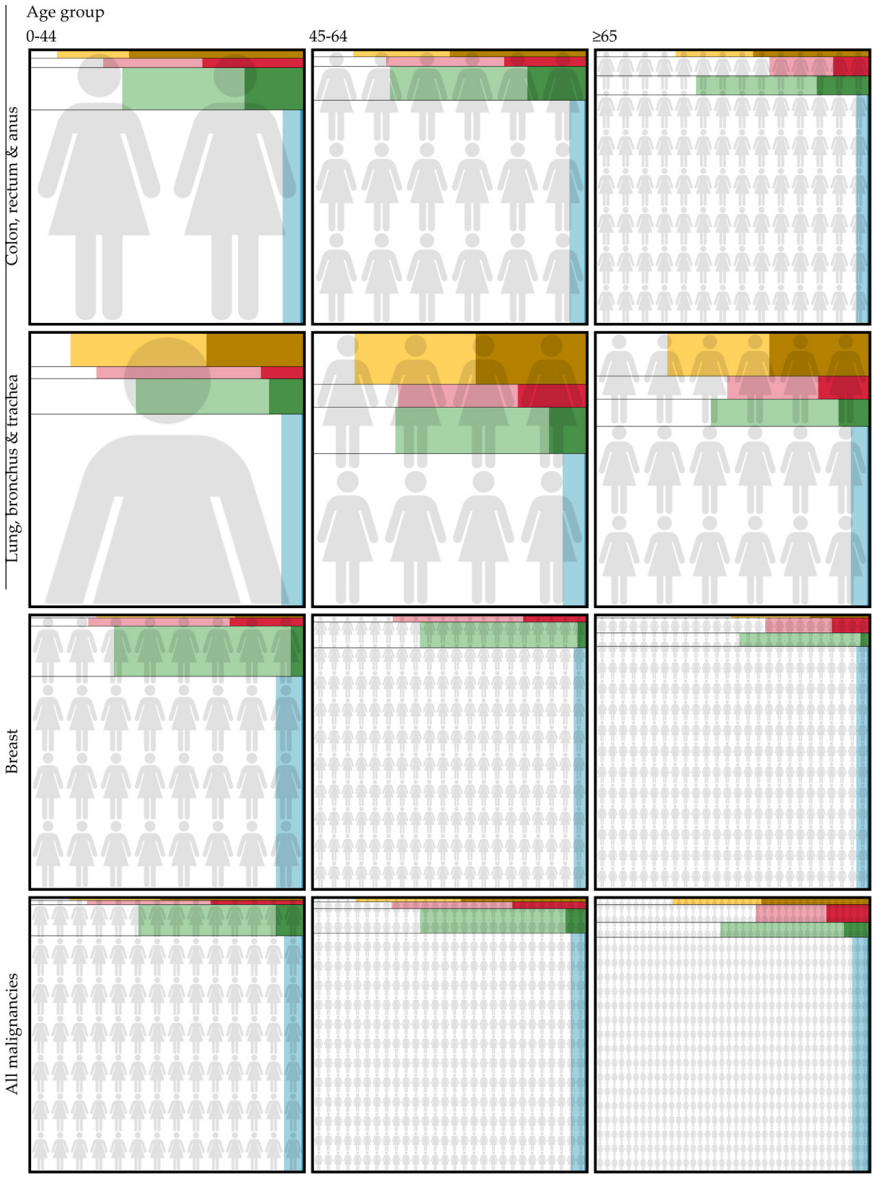


Fig. 4 – Female cancer survivors, UK, 2008; by age, type of cancer and survivorship state. In each tile, the rows from top to bottom correspond to the 4 temporal phases of survivorship (A-D) respectively, with the vertical bands from left to right corresponding to states of increasing acute health service utilisation (see Fig. 2 and Box 1 for definitions). Each female icon represents approximately 1,000 cancer survivors (exact counts are in Table 1). The proportion of survivors in temporal phase

D (more than 1 year from both diagnosis and death) and with a 'high' level of acute health service utilisation (the rightmost area of the bottom row) may be too small to be visible. Similarly, the proportion of breast cancer survivors in temporal phase A (within one year of both diagnosis and death – the top row) may be too small to be visible.

Table 1 – Number of survivors in the United Kingdom, 2008. By temporal phase of survivorship[a] and intensity of cancer related hospitalisation[b].

Sex			Age group											
			0–44 Intensity of hospitalisation				45–64 Intensity of hospitalisation				⩾65 Intensity of hospitalisation			
	Type of Cancer	Temporal phase	None	Low	High	Total	None	Low	High	Total	None	Low	High	Total
Male														
	Colon, rectum and anus	A	<10	20	30	60	80	260	270	610	530	710	840	2,080
		B	10	40	30	80	290	510	270	1,060	4,120	2,130	1,070	7,320
		C	130	150	80	360	1,050	1,820	790	3,670	2,590	3,650	1,590	7,840
		D	1,490	90	<10	1,590	18,760	1,430	160	20,350	72,400	4,700	410	77,520
		Total	1,640	310	150	2,090	20,180	4,010	1,500	25,690	79,640	11,200	3,910	94,750
	Lung, bronchus and trachea	A	<10	30	30	70	210	660	520	1,400	830	1,420	1,260	3,500
		B	<10	<10	<10	20	180	260	120	550	930	770	370	2,070
		C	20	40	<10	70	320	630	130	1,090	830	1,090	220	2,130
		D	280	10	<10	290	3,310	270	20	3,600	22,360	1,500	90	23,950
		Total	310	90	50	440	4,030	1,820	790	6,640	24,950	4,780	1,940	31,660
	Prostate	A	<10	<10	<10	<10	80	130	70	280	1,110	890	740	2,740
		B	<10	<10	<10	<10	220	350	180	750	7,800	5,860	3,210	16,880
		C	20	10	<10	40	5,020	3,590	140	8,740	17,190	8,370	600	26,160
		D	120	10	<10	130	22,830	1,830	80	24,740	152,850	18,960	1,150	172,970
		Total	140	30	10	180	28,150	5,890	470	34,510	178,960	34,080	5,700	218,740
	All malignant neoplasms	A	80	210	360	650	880	2,260	2,500	5,640	4,180	6,170	6,680	17,030
		B	220	360	300	880	1,960	2,680	1,790	6,430	21,210	14,480	8,060	43,750
		C	2,400	2,410	1,150	5,950	10,330	10,810	2,740	23,880	24,500	19,820	4,270	48,600
		D	59,200	4,020	560	63,780	142,560	12,260	1,150	155,970	398,000	45,800	2,840	446,640
		Total	61,890	6,990	2,370	71,250	155,730	28,010	8,180	191,920	447,890	86,270	21,850	556,020
Female														
	Colon, rectum and anus	A	<10	10	30	50	50	140	190	390	510	500	750	1,750
		B	20	30	30	80	180	300	200	680	4,010	1,470	820	6,300
		C	110	150	70	330	710	1,260	540	2,510	2,310	2,790	1,190	6,290
		D	1,550	110	10	1,680	15,140	910	100	16,150	73,780	3,010	310	77,090
		Total	1,690	300	150	2,130	16,080	2,610	1,030	19,720	80,600	7,760	3,070	91,430
	Lung, bronchus and trachea	A	<10	30	20	60	160	480	440	1,080	750	1,060	1,030	2,840
		B	<10	10	<10	20	150	220	120	500	760	520	290	1,570
		C	30	30	<10	70	300	570	130	1,000	780	860	200	1,850
		D	340	30	<10	370	3,040	260	20	3,320	11,320	740	40	12,090
		Total	380	110	40	530	3,660	1,530	710	5,900	13,600	3,180	1,560	18,350
	Breast	A	30	60	30	120	130	250	140	520	830	480	360	1,670
		B	150	370	190	710	1,260	2,110	1,020	4,390	11,200	4,390	2,400	17,990
		C	1,430	3,060	210	4,700	7,610	11,430	610	19,660	8,250	7,020	500	15,770
		D	17,950	1,870	80	19,890	174,890	8,240	380	183,510	268,030	11,200	840	280,070
		Total	19,560	5,360	510	25,430	183,890	22,030	2,150	208,080	288,300	23,090	4,100	315,490
	All malignant neoplasms	A	100	230	360	690	670	1,650	1,970	4,290	3,570	4,100	5,010	12,680
		B	310	680	510	1,490	2,690	4,230	2,550	9,470	25,610	11,570	6,890	44,070
		C	4,390	5,620	1,140	11,160	13,460	18,490	2,560	34,510	17,200	17,270	3,510	37,970
		D	76,500	5,240	540	82,280	318,620	17,370	1,200	337,190	573,070	32,000	2,460	607,520
		Total	81,300	11,770	2,550	95,620	335,440	41,730	8,280	385,460	619,440	64,940	17,870	702,250

Numbers may not sum to group totals since all are rounded to the nearest 10. Numbers less than 10 are suppressed.

[a] Phase A = less than 1 year from death and less than 1 year from diagnosis; phase B = less than 1 year from death and more than 1 year from diagnosis; phase C = more than 1 year from death and less than 1 year from diagnosis; phase D = more than 1 year from death and more than 1 year from diagnosis.

[b] None = no time is spent as an admitted hospital patient; low = time spent as an admitted hospital patient accounts for some, but no more than 10% of, person time; high = time spent as an admitted hospital patient accounts for more than 10% of person time.

Table 2 – Percentage of survivors in the UK, 2008. By temporal phase of survivorship[a] and intensity of cancer related hospitalisation[b].

Sex			Age group											
			0–44 Intensity of hospitalisation				45–64 Intensity of hospitalisation				⩾65 Intensity of hospitalisation			
	Type of Cancer	Temporal phase	None	Low	High	Total	None	Low	High	Total	None	Low	High	Total
Male														
	Colon, rectum and anus	A	0.4	1.2	1.5	3.1	0.3	1.0	1.1	2.4	0.6	0.8	0.9	2.2
		B	0.7	1.9	1.2	3.8	1.1	2.0	1.1	4.1	4.3	2.2	1.1	7.7
		C	6.2	7.3	3.8	17.3	4.1	7.1	3.1	14.3	2.7	3.9	1.7	8.3
		D	71.1	4.3	0.5	75.8	73.0	5.5	0.6	79.2	76.4	5.0	0.4	81.8
		Total	78.3	14.7	7.0	100.0	78.6	15.6	5.8	100.0	84.0	11.8	4.1	100.0
	Lung, bronchus and trachea	A	1.7	7.6	7.0	16.3	3.2	10.0	7.9	21.1	2.6	4.5	4.0	11.1
		B	0.7	1.4	1.5	3.6	2.7	3.9	1.8	8.4	2.9	2.4	1.2	6.5
		C	5.0	7.9	2.1	15.0	4.9	9.5	2.0	16.4	2.6	3.4	0.7	6.7
		D	62.4	2.6	0.2	65.1	49.9	4.0	0.2	54.2	70.6	4.7	0.3	75.7
		Total	69.8	19.5	10.7	100.0	60.7	27.5	11.9	100.0	78.8	15.1	6.1	100.0
	Prostate	A	0.0	0.5	1.0	1.6	0.2	0.4	0.2	0.8	0.5	0.4	0.3	1.3
		B	0.8	0.5	1.3	2.6	0.6	1.0	0.5	2.2	3.6	2.7	1.5	7.7
		C	13.2	7.3	2.4	22.9	14.5	10.4	0.4	25.3	7.9	3.8	0.3	12.0
		D	64.5	6.8	1.7	73.0	66.2	5.3	0.2	71.7	69.9	8.7	0.5	79.1
		Total	78.5	15.2	6.4	100.0	81.6	17.1	1.4	100.0	81.8	15.6	2.6	100.0
	All malignant neoplasms	A	0.1	0.3	0.5	0.9	0.5	1.2	1.3	2.9	0.8	1.1	1.2	3.1
		B	0.3	0.5	0.4	1.2	1.0	1.4	0.9	3.4	3.8	2.6	1.4	7.9
		C	3.4	3.4	1.6	8.3	5.4	5.6	1.4	12.4	4.4	3.6	0.8	8.7
		D	83.1	5.6	0.8	89.5	74.3	6.4	0.6	81.3	71.6	8.2	0.5	80.3
		Total	86.9	9.8	3.3	100.0	81.1	14.6	4.3	100.0	80.6	15.5	3.9	100.0
Female														
	Colon, rectum and anus	A	0.2	0.6	1.5	2.4	0.3	0.7	1.0	2.0	0.6	0.5	0.8	1.9
		B	1.0	1.3	1.4	3.7	0.9	1.5	1.0	3.5	4.4	1.6	0.9	6.9
		C	5.2	7.0	3.4	15.5	3.6	6.4	2.7	12.7	2.5	3.0	1.3	6.9
		D	72.6	5.2	0.6	78.5	76.8	4.6	0.5	81.9	80.7	3.3	0.3	84.3
		Total	79.0	14.1	6.9	100.0	81.5	13.2	5.2	100.0	88.2	8.5	3.4	100.0
	Lung, bronchus and trachea	A	1.8	6.1	4.3	12.2	2.8	8.2	7.4	18.4	4.1	5.8	5.6	15.5
		B	1.1	2.7	0.7	4.5	2.6	3.7	2.1	8.4	4.1	2.9	1.6	8.6
		C	5.1	6.4	1.6	13.1	5.2	9.6	2.3	17.0	4.2	4.7	1.1	10.1
		D	64.7	5.2	0.3	70.2	51.4	4.4	0.3	56.2	61.7	4.0	0.2	65.9
		Total	72.6	20.4	7.0	100.0	62.0	25.9	12.1	100.0	74.1	17.3	8.5	100.0
	Breast	A	0.1	0.2	0.1	0.5	0.1	0.1	0.1	0.2	0.3	0.2	0.1	0.5
		B	0.6	1.5	0.8	2.8	0.6	1.0	0.5	2.1	3.5	1.4	0.8	5.7
		C	5.6	12.0	0.8	18.5	3.7	5.5	0.3	9.4	2.6	2.2	0.2	5.0
		D	70.6	7.3	0.3	78.2	84.1	4.0	0.2	88.2	85.0	3.5	0.3	88.8
		Total	76.9	21.1	2.0	100.0	88.4	10.6	1.0	100.0	91.4	7.3	1.3	100.0
	All malignant neoplasms	A	0.1	0.2	0.4	0.7	0.2	0.4	0.5	1.1	0.5	0.6	0.7	1.8
		B	0.3	0.7	0.5	1.6	0.7	1.1	0.7	2.5	3.6	1.6	1.0	6.3
		C	4.6	5.9	1.2	11.7	3.5	4.8	0.7	9.0	2.4	2.5	0.5	5.4
		D	80.0	5.5	0.6	86.0	82.7	4.5	0.3	87.5	81.6	4.6	0.3	86.5
		Total	85.0	12.3	2.7	100.0	87.0	10.8	2.1	100.0	88.2	9.2	2.5	100.0

Percentages sum to 100% for each type of cancer/age group (subject to rounding errors).

[a] Phase A = less than 1 year from death and less than 1 year from diagnosis; phase B = less than 1 year from death and more than 1 year from diagnosis; phase C = more than 1 year from death and less than 1 year from diagnosis; phase D = more than 1 year from death and more than 1 year from diagnosis.

[b] None = no time is spent as an admitted hospital patient; low = time spent as an admitted hospital patient accounts for some, but no more than 10% of, person time; high = time spent as an admitted hospital patient accounts for more than 10% of person time.

At the end of 2008, approximately 81% of all male survivors and 87% of all female survivors were more than 1 year from diagnosis and more than 1 year from death. There was little variation in these proportions by age, except for male survivors aged under 45 years for whom this proportion rose to 90%. The proportion of lung cancer survivors who were more than 1 year from both diagnosis and death was markedly lower (72% of male survivors and 64% of female survivors). There was little acute health service utilisation among survivors in this phase, regardless of age or type of cancer. Overall, 1.57 million cancer survivors in the UK at the end of 2008 (78%) were more than 1 year from both diagnosis and death, and had no cancer related hospital admissions.

240,000 of the two million cancer survivors in the UK (14.8% of male survivors and 10.0% of female survivors) were in the 'low' acute health care utilisation category. Sixty one thousand (4.0% of male survivors and 2.4% of female survivors) were in the 'high' category. The proportion of lung cancer survivors in the 'high' category was much larger (7.2% of males and 9.3% of females), particularly in the age group 45–64 years where the proportion rose to around 12%. Conversely, only 1.2% of female breast cancer survivors had a 'high' level of acute health care utilisation.

147,000 cancer survivors were in the last year of their life (9.0% of male survivors and 6.2% of female survivors). It was these survivors who had the highest intensity of cancer related acute health service utilisation, particularly those who were also in the first year since diagnosis. Forty one percent of the 41,000 survivors who were less than 1 year from both diagnosis and death had a high intensity of acute health service utilisation, compared with 19% of the 106,000 survivors who were in the last year of their life but more than 1 year from diagnosis.

4. Discussion

The linked dataset allowed an analysis of all recorded episodes of hospital in-patient or day case healthcare among registered cancer survivors in England during 2006. A person-time approach allowed us to quantify the intensity of hospitalisation, and the associated burden on cancer survivors and the health service, by considering the amount of time spent in hospital, rather than just the number of admissions.

The hospital activity data featured details of all types of in-patient and day case care delivered to cancer survivors, not just that which related to cancer. Each record had up to 14 diagnostic codes (using the ICD-10 classification) and 12 operation procedure codes, with the first of each of these codes intended to indicate the primary diagnosis or intent of the treatment. For simplicity, we categorised episodes as 'related to cancer' or 'not related to cancer' according to the diagnostic codes only. This was an intentionally broad categorisation, designed to take account of the wide range of health problems associated with cancer and the side effects of its treatment. Nonetheless the specificity of the definition is unproven and it is not possible to say exactly which kinds of treatment following a cancer diagnosis are included. However, an analysis of the complementary 'non-cancer related' care (not presented here) indicates that the cancer related definition does achieve its goal by removing some of the background hospitalisation experienced by this population. The proportion of time spent by cancer survivors in hospital for non-cancer related care was much lower than for cancer related care, and was generally constant regardless of time since diagnosis. In this analysis, therefore, 'cancer related' hospitalisation may be considered as that directly or indirectly caused by, or in some way associated with, a cancer diagnosis.

Some survivors will be diagnosed with additional primary cancers some time after their first diagnosis. Indeed, many studies have shown elevated cancer incidence rates in those previously diagnosed with cancer compared with the general population[17,18] and the prevalence of multiple malignancies among cancer survivors has been shown to be around 7%.[19] However, treatment received for subsequent cancers is indistinguishable in this analysis from that received for the initial cancer, a fact to be born in mind when considering the 'time since diagnosis' dimension of this study.

A small number of survivors in the cancer registry data will have had a diagnosis pre-1990 as well as in the period 1990–2006. However, since details of diagnoses made before 1990 were not available in the linked dataset, it was necessary to assume that all survivors alive during 2006 had their first cancer diagnosis in the period 1990–2006. The effect of this assumption is a possible small re-distribution of survivors between the tumour sites studied, but is not considered to be a significant limitation.

Survivorship states were defined according to time since diagnosis, time until death, and proportion of time spent admitted to hospital. The first year following diagnosis was considered to be potentially important, since this is the time during which cancer patients receive initial treatment, the success of which may significantly affect their subsequent health and well-being. The final year before death was also considered to be potentially important since for many people who die from cancer there is a period of health deterioration in the months beforehand. For some survivors post-diagnosis survival is short – a separate temporal phase was therefore defined by the intersection of the first year following diagnosis and the last year of life. The period of survivorship which is more than 1 year from both diagnosis and death also defined a separate temporal phase and may be characterised by periods of remission, relapse, disease monitoring, and/or eventual 'cure'.

Cancer prevalence is driven by cancer incidence and survival, and the distribution of survivors between temporal phases of survivorship can largely be explained by incidence and survival characteristics. For example, lung cancer has a universally poor prognosis (age-standardised five-year relative survival rates in England and Wales are under 10%[20]) and accordingly a large proportion of lung cancer survivors were less than 1 year from both diagnosis and death. On the other hand, one quarter of prostate cancer survivors aged 45–64 were less than 1 year from diagnosis but not less than 1 year from death, reflecting the relatively good prognosis of this disease and the rapidly increasing incidence rates brought about by the diagnostic use of the PSA test since the early 1990s.[21]

In recent decades, the number of cancer survivors in the UK has grown steadily each year.[1] The distribution of survivors between temporal phases, and the intensity of acute health service utilisation within them, provides an insight into what is meant by the term 'cancer survivor', especially given the current national survivorship initiatives in the UK and the movement towards understanding cancer as a chronic illness. For example, the majority of UK cancer survivors (1.69 million of the 2.00 million) are more than 1 year from both diagnosis and death, and the degree of acute health service utilisation in this phase is small – 1.57 million are in a period characterised by no cancer related acute health service utilisation and these survivors account for 78% of all UK can-

cer survivors. Cancer survivors can now realistically expect to live longer, but this analysis suggests that the primary burden of cancer on the health service still comes from survivors in the first year following diagnosis and/or near the end of their life. The term 'cancer survivor' was originally proposed by the US National Coalition for Cancer Survivorship in 1986 at a time when "cancer was a disease that people needed to learn to fight" but has, according to some, become "so muddy that...a new definition is needed".[22] Alternative terms include people 'living with or beyond cancer' or 'cancer patients'. However, this analysis has shown that survivorship is heterogeneous, and finding a single term to usefully define everyone who has ever been diagnosed with cancer may not be possible.

A limitation of this work is that it only considered *admitted* hospital episodes (in-patients and day cases). Visits to general practitioner surgeries and other outpatient clinics are not captured in this analysis, but much of the observation and monitoring of survivors (especially those who are more than 1 year from both diagnosis and death) is carried out in outpatient clinics.[23] Neither does this analysis consider the personal psychosocial or general health burden of cancer on survivors – the trauma of being diagnosed with a life-threatening illness such as cancer is associated with post traumatic stress disorder, depression, and other mental disorders,[24,25] and cancer survivors have been found to have poorer general health outcomes than individuals who have not been diagnosed with cancer.[26] Cancer survivors are also likely to face day-to-day struggles (such as financial, emotional, relationship, and employment difficulties) even if they have no need for treatment in hospital.[27,28] These issues present significant burdens to cancer survivors and should be kept in mind when considering the distribution of survivors between the states defined in this analysis.

The extent to which the population of cancer survivors in the UK is receiving care and treatment in hospital is central to understanding the burden of cancer on society. The findings contained here will be of interest to health service providers keen to quantify the volume of acute health care administered to cancer survivors, and the associated financial burden, as well as to survivors themselves.

Conflict of interest statement

None declared.

Acknowledgements

The study design and concepts were developed jointly by Jacob Maddams, Henrik Møller, and Martin Utley. Jacob Maddams conducted the analysis and wrote the first manuscript draft. Henrik Møller and Martin Utley provided critical revisions of the manuscript. All authors have approved the final version.

The corresponding author, Jacob Maddams, had full access to all the data in this study and had final responsibility for the decision to submit the manuscript for publication.

This research was funded by Macmillan Cancer Support, who had no involvement in the study design, the collection, analysis, and interpretation of the data, in the writing of the report, or in the decision to submit this paper for publication.

This paper is a contribution from the National Cancer Intelligence Network and is based on the information collected and quality assured by the regional cancer registries in England (http://www.ukacr.org; http://www.ncin.org.uk). The Thames Cancer Registry in King's College London receives funding from the Department of Health for England. The views expressed in the article are those of the authors and not necessarily those of the Department of Health. The Clinical Operational Research Unit receives funding from the Department of Health for England Policy Research Programme.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ejca.2011.04.015.

REFERENCES

1. Maddams J, Brewster D, Gavin A, et al. Cancer prevalence in the United Kingdom: estimates for 2008. *Br J Cancer* 2009;**101**(3):541–7.
2. UK Government Department of Health Cancer Reform Strategy. [Accessed March 2, 2011]; http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_081006.
3. Maliski SL, Heilemann MV, McCorkle R. From "death sentence" to "good cancer": couples' transformation of a prostate cancer diagnosis. *Nurs Res* 2002;**51**(6):391–7.
4. Hubbard G. Cancer: no longer a death sentence. *Br J Community Nurs* 2010;**15**(5):213.
5. Markman M. Viewing ovarian cancer as a "chronic disease": What exactly does this mean? *Gynecol Oncol* 2006;**100**(2):229–30.
6. Yabroff KR, Lamont EB, Mariotto A, et al. Cost of care for elderly cancer patients in the United States. *J Natl Cancer Inst* 2008;**100**(9):630–41.
7. Parkin DM, Bray F, Ferlay J, Pisani P. Estimating the world cancer burden: globocan 2000. *Int J Cancer* 2001;**94**(2):153–6.
8. Pisani P, Bray F, Parkin DM. Estimates of the world-wide prevalence of cancer for 25 sites in the adult population. *Int J Cancer* 2002;**97**(1):72–81.
9. Benhamiche-Bouvier AM, Clinard F, Phelip JM, Rassiat E, Faivre J. Colorectal cancer prevalence in France. *Eur J Cancer Prev* 2000;**9**(5):303–7.
10. Gatta G, Capocaccia R, Berrino F, et al. Colon cancer prevalence and estimation of differing care needs of colon cancer patients. *Ann Oncol* 2004;**15**(7):1136–42.
11. Louchini R, Beaupre M, Demers AA, Goggin P, Bouchard C. Trends in cancer prevalence in Quebec. *Chronic Dis Can* 2006;**27**(3):110–9.
12. Capocaccia R, De Angelis R, Frova L, et al. Estimation and projections of colorectal cancer trends in Italy. *Int J Epidemiol* 1997;**26**(5):924–32.
13. McCarthy M. New measure of cancer prevalence may aid health planning. *Lancet Oncol* 2002;**3**(2):70.
14. HES Online: Hospital Episode Statistics. [Accessed March 2, 2011]; http://www.hesonline.nhs.uk/Ease/servlet/ContentServer?siteID=1937.

15. Carstensen B. Examples and programs for splitting follow-up time in cohort studies. [Accessed March 2, 2011]; http://www.staff.pubhealth.ku.dk/~bxc/Lexis/.
16. Biography of Piet Mondrian. [Accessed March 2, 2011]; http://www.tate.org.uk/servlet/ArtistWorks?cgroupid=999999961&artistid=1651&page=1&sole=y&collab=y&attr=y&sort=default&tabview=bio.
17. Evans HS, Lewis CM, Robinson D, et al. Incidence of multiple primary cancers in a cohort of women diagnosed with breast cancer in southeast England. *Br J Cancer* 2001;**84**(3):435–40.
18. Evans HS, Moller H, Robinson D, et al. The risk of subsequent primary cancers after colorectal cancer in southeast England. *Gut* 2002;**50**(5):647–52.
19. Liu L, de Vries E, Louwman M, et al. Prevalence of multiple malignancies in the Netherlands in 2007. *Int J Cancer* 2011;**128**(7):1659–67.
20. Lung cancer survival statistics. [Accessed March 2, 2011]; http://www.info.cancerresearchuk.org/cancerstats/types/lung/survival/.
21. Evans HS, Moller H. Recent trends in prostate cancer incidence and mortality in southeast England. *Eur Urol* 2003;**43**(4):337–41.
22. Twombly R. What's in a name: who is a cancer survivor? *J Natl Cancer Inst* 2004;**96**(19):1414–5.
23. Khan NF, Carpenter L, Watson E, Rose PW. Cancer screening and preventative care among long-term cancer survivors in the United Kingdom. *Br J Cancer* 2010;**102**(7):1085–90.
24. Gregurek R, Bras M, Dordevic V, Ratkovic AS, Brajkovic L. Psychological problems of patients with cancer. *Psychiatr Danub* 2010;**22**(2):227–30.
25. Smith MY, Redd WH, Peyser C, Vogl D. Post-traumatic stress disorder in cancer: a review. *Psychooncology* 1999;**8**(6):521–37.
26. Yabroff KR, Lawrence WF, Clauser S, Davis WW, Brown ML. Burden of illness in cancer survivors: findings from a population-based national sample. *J Natl Cancer Inst* 2004;**96**(17):1322–30.
27. Foster C, Wright D, Hill H, Hopkinson J, Roffe L. Psychosocial implications of living 5 years or more following a cancer diagnosis: a systematic review of the research evidence. *Eur J Cancer Care (Engl)* 2009;**18**(3):223–47.
28. Short PF, Vasey JJ, BeLue R. Work disability associated with cancer survivorship and other chronic conditions. *Psychooncology* 2008;**17**(1):91–7.

British Journal of Cancer (2011) 105, S38 – S45

www.bjcancer.com

Full Paper

A person-time analysis of hospital activity among cancer survivors in England

J Maddams*[,1], M Utley[2] and H Møller[1]

[1]*King's College London, Thames Cancer Registry, 1st Floor, Capital House, 42 Weston Street, London SE1 3QD, UK;* [2]*Clinical Operational Research Unit, University College London, 4 Taviton Street, London WC1H 0BT, UK*

BACKGROUND: There are around 2 million cancer survivors in the UK. This study describes the inpatient and day case hospital activity among the population of cancer survivors in England. This is one measure of the burden of cancer on the individual and the health service.

METHODS: The national cancer registry data set for England (1990–2006) is linked to the NHS Hospital Episode Statistics (HES) database. Cohorts of survivors were defined as those people recorded in the cancer registry data with a diagnosis of breast, colorectal, lung or prostate cancer before 2007. The person-time of prevalence in 2006 for each cohort of survivors was calculated according to the cancer type, sex, age and time since diagnosis. The corresponding HES episodes of care in 2006 were used to calculate the person-time of admitted hospital care for each cohort of survivors. The average proportion of time spent in hospital by survivors in each cohort was calculated as the summed person-time of hospital activity divided by the summed person-time of prevalence. The analysis was conducted separately for cancer-related episodes and non-cancer-related episodes.

RESULTS: Lung cancer survivors had the highest intensity of cancer-related hospital activity. For all cancers, cancer-related hospital activity was highest in the first year following diagnosis. Breast and prostate cancer survivors had peaks of cancer-related hospital activity in the relatively young and relatively old age groups. The proportion of time spent in hospital for non-cancer-related care was much lower than that for cancer-related care and increased gradually with age but was generally constant regardless of time since diagnosis.

CONCLUSION: The person-time approach used in this study is more revealing than a simple enumeration of cancer survivors and hospital admissions. Hospital activity among cancer survivors is highest soon after diagnosis. The effect of age on the amount of hospital activity is different for each type of cancer.

British Journal of Cancer (2011) **105,** S38 – S45; doi:10.1038/bjc.2011.421 www.bjcancer.com

Keywords: cancer burden; cancer survivorship; health-service use

Cancer survivors are defined as people who are alive following a diagnosis of cancer from some point in their past. In UK, there are approximately 2 million cancer survivors (approximately 1.66 million of these in England), and this number is increasing by 3% per year (Maddams *et al*, 2009). The size and demographics of the population of cancer survivors has been described, but there is still a need for more detailed analyses of the burden of cancer on the individual and on health-service resources.

This work presents a person-time analysis of a linked cancer registry and health-service activity data set. 'Person-time of survivorship' refers to the total time at risk of hospitalisation experienced by a population of survivors, and 'person-time of hospitalisation' refers to the total time spent in hospital by this population. Acute health-service utilisation among cancer survivors in England is described according to age and time since diagnosis, for both cancer- and non-cancer-related care. The work presented here is complementary to that previously published by the same authors (Maddams *et al*, 2011).

*Correspondence: Dr J Maddams; E-mail: jacob.maddams@kcl.ac.uk

MATERIALS AND METHODS

Data

The analysis was based on two linked data sets (Maddams *et al*, 2011). The first was a national merged cancer registry data set featuring patient and diagnostic information relating to all cancers recorded by the eight regional population-based cancer registries in England. The merging process was carried out by the staff at the Thames Cancer Registry (TCR), on behalf of the United Kingdom Association of Cancer Registries (UKACR). This provided 100% coverage of geographical regions of England and included all registered cancers diagnosed from 1990 to 2006 (inclusive).

This data set was linked to an extract from the NHS Hospital Episode Statistics (HES). Hospital Episode Statistics is a record-level data repository managed by the National Health Service Information Centre on behalf of the Secretary of State for Health. It contains patient, clinical and administrative details for admitted patients and outpatients treated in NHS hospitals in England, and is mainly populated by extracts from routine data flows exchanged between health-care providers and commissioners (NHS Health and Social Care Information Centre, 2010). Each HES record

defines a finished consultant episode (FCE) of care under a given consultant in a given NHS provider. A patient's journey from admission to discharge may be made up of several FCEs. The English Cancer Registries' National HES extract is a subset of the complete HES database and contains only episodes for admitted patients (i.e., inpatients and day patients) who have at least one recorded episode 'for or with' cancer. An episode is considered to be 'for or with' cancer if any of its 14 diagnostic fields contain an ICD-10 code between C00 and C97 (malignant neoplasms), between D01 and D48 (*in situ*, benign or uncertain neoplasms) or equal to O01 (hydatidiform mole).

The linkage between these two data sets was designed and developed jointly by the TCR and the Northern and Yorkshire Cancer Registry and Information Service on behalf of the UKACR. The methodology matches fields such as sex, date of birth, date of death, NHS number and postcode across the two data sets and provides a unique patient identifier for matched patients. Using the patient identifier as a link, it is possible to extract all HES episodes of care for a cohort of cancer patients, as defined in the national cancer registry data set.

The national cancer registry data set was used to identify a cohort of cancer survivors who had been both diagnosed with a malignant neoplasm (ICD-10 C00–C97 excluding C44) in the period 1990–2006, and alive for at least some portion of 2006 (Maddams *et al*, 2011). Subcohorts were defined according to the type of cancer diagnosis received: colon, rectum and anus cancers (ICD-10 C18–C21); lung, bronchus and trachea cancers (ICD-10 C33–C34); prostate cancer (ICD-10 C61); and female breast cancer (ICD-10 C50). Survivors with multiple diagnoses entered multiple subcohorts. The unique patient identifiers were then used to extract all the HES episodes of care that occurred in, or overlapped, the year 2006 for the cohorts of cancer survivors. As the English Cancer Registries' HES extract contains all inpatient and day case episodes for those patients with at least one cancer-related episode in the entire HES data set, the extracted cancer survivor episodes may or may not mention cancer. Episodes not mentioning cancer may have pertained to an entirely unrelated condition, or (less likely) may have been incorrectly coded. Some cancer survivors had no matching HES episode of care in 2006, because of a failure to register episodes that did occur, a failure in the matching procedure between the two data sets or simply because they were not admitted to hospital in 2006. In this analysis, matched episodes of care were considered to be 'cancer related' if one of the 14 diagnostic codes was between C00 and C97 (excluding C44); otherwise, they were considered to be 'non-cancer related'.

Hospital activity among cancer survivors

A person-time approach was used to describe the effect of age and time since diagnosis on the amount of hospital activity in the population of cancer survivors. When considering the interaction of survivors with the health service, a person-time approach has a distinct advantage, as it allows us to estimate the proportion of survivor time spent admitted to hospital, as well as to count the number of admissions. The calendar year 2006 was the period of analysis. The effect of age was analysed in 5-year age groups, and the effect of time since diagnosis was analysed in 1-year periods and broader periods (<1, 1–5 and 5–16 years since diagnosis).

Diagnoses were available for the period 1990–2006, and thus there were some survivors in the 2006 cohort who had been diagnosed more than 16 years (but less than 17 years) previously. However, the cohort did not contain all such survivors (as no diagnoses from 1989 were available), and therefore the maximum time since diagnosis considered was 16 years.

For each survivor in each cohort, the person-time for which they were prevalent in 2006 (i.e., that which was postdiagnosis and predeath and overlapped the calendar year 2006) was split into segments according to the time points at which the indexing variables 'time since diagnosis' and 'age' changed (Maddams *et al*, 2011). The amount of time spent in hospital by cancer survivors was counted in a similar way, according to age and time since diagnosis, for cancer-related and non-cancer-related episodes of care separately. This was achieved using the SAS software package (SAS Institute Inc., Cary, NC, USA) and code developed by JM specifically for this task, similar to the program Lexis.sas by Carstensen (2007).

The proportion of person-time each cohort of survivors spent admitted to hospital, according to age and time since diagnosis, was then calculated by dividing the total population person-time spent in HES activity by the total population person-time of prevalence. This quantity may be interpreted as the mean proportion of time spent in HES activity by a survivor of a given age and time since diagnosis sampled at random. An alternative interpretation is that the quantity provides an estimate of the probability of that survivor being an admitted hospital patient on any given day in the period. It may be presented either as the number of days spent admitted to hospital per 100 person-days or as a percentage of time. Confidence intervals (CIs) for this proportion were calculated by using a standard binomial model, modified to include a dependence parameter (Klotz, 1973; Budescu, 1985), to account for the fact that days spent in hospital for an individual survivor are likely to be clustered together in episodes or spells of care, rather than randomly distributed in time. The dependence was assumed to be of the first-order Markov type, and both the dependence parameters and CIs were estimated using the methods described in Budescu (1985).

RESULTS

Tables 1 and 2, and the corresponding Figures 1 and 2, display the average number of days (per 100 person-days) that cancer survivors spent admitted to hospital in England during 2006, by 5-year age group or 1-year time since diagnosis period. Data are presented for cancer-related and non-cancer-related episodes of care separately, but are omitted if less than 20 person-years of survivorship contributed to the estimate.

In all age groups older than 35 years, the number of days admitted to hospital for cancer-related episodes of care (per 100 person-days) was highest among lung cancer survivors; it was lowest among prostate and breast cancer survivors. Among prostate cancer survivors, the relatively young and relatively old age groups spent a higher number of days admitted to hospital for cancer-related episodes (per 100 person-days) than the middle-age groups (55–75 years). A similar trend was observed among breast cancer survivors, although the observed increase in the number of days spent admitted to hospital above the age of 75 years was not as significant as among prostate cancer survivors. Among male lung and colorectal cancer survivors, the number of days spent admitted to hospital for cancer-related episodes of care (per 100 person-days) generally decreased as age increased – this association was reversed in male lung cancer survivors under the age of 50 years, but CIs were wide because of a relatively small number of lung cancer survivors in these age groups. Similar age effects were observed among female lung and colorectal cancer survivors.

The number of days spent admitted to hospital for non-cancer-related episodes of care (per 100 person-days) generally increased as age increased above 60 years. Colorectal and lung cancer survivors spent a similar number of days admitted for non-cancer-related episodes of care (per 100 person-days); however, prostate and breast cancer survivors spent significantly fewer.

The number of days survivors spent admitted to hospital (per 100 person-days) for cancer-related episodes of care was much higher in the first year following diagnosis than at any other time. In the first 5 years after diagnosis, it was highest among survivors of lung cancer.

Table 1 Mean admitted patient hospital activity (cancer related and non-cancer related) among cancer survivors, England, 2006. Number of days in hospital per 100 person-days,[¶] by 5-year age group

	Cancer diagnosis							
	Colon, rectum and anus		Lung, bronchus and trachea		Breast		Prostate	
	Cancer related	Non-cancer related	Cancer related	Non-cancer related	Cancer related	Non-cancer related	Cancer related	Non-cancer related
Males								
Age group (in years)								
0–4								
5–9								
10–14								
15–19								
20–24	3.4 (2.5–4.2)	1.0 (0.5–1.5)						
25–29	3.2 (2.6–3.8)	0.9 (0.7–1.2)	2.3 (0.8–3.9)	0.4 (0.0–0.9)				
30–34	2.4 (2.0–2.9)	0.3 (0.2–0.5)	2.0 (1.1–2.9)	0.1 (0.0–0.3)				
35–39	2.5 (2.2–2.8)	0.5 (0.4–0.6)	3.6 (2.9–4.3)	0.4 (0.0–0.7)				
40–44	2.0 (1.8–2.2)	0.5 (0.4–0.5)	3.9 (3.5–4.4)	0.3 (0.1–0.5)			1.8 (1.2–2.4)	0.2 (0.0–0.5)
45–49	2.1 (2.0–2.2)	0.6 (0.6–0.7)	4.6 (4.3–5.0)	0.4 (0.3–0.6)			1.2 (0.9–1.4)	0.2 (0.1–0.3)
50–54	2.2 (2.1–2.3)	0.5 (0.5–0.6)	4.4 (4.2–4.6)	0.6 (0.5–0.7)			0.9 (0.8–1.0)	0.2 (0.2–0.3)
55–59	2.1 (2.0–2.2)	0.5 (0.5–0.6)	4.7 (4.5–4.8)	0.6 (0.5–0.7)			0.6 (0.6–0.7)	0.2 (0.2–0.2)
60–64	1.9 (1.9–2.0)	0.7 (0.6–0.7)	4.2 (4.0–4.3)	0.7 (0.6–0.8)			0.6 (0.6–0.7)	0.3 (0.3–0.3)
65–69	2.0 (1.9–2.0)	0.7 (0.7–0.7)	4.1 (4.0–4.2)	0.6 (0.6–0.7)			0.6 (0.6–0.7)	0.3 (0.3–0.3)
70–74	1.7 (1.7–1.8)	0.8 (0.8–0.9)	4.0 (3.9–4.1)	0.7 (0.7–0.8)			0.7 (0.7–0.7)	0.4 (0.4–0.4)
75–79	1.7 (1.6–1.7)	0.9 (0.9–1.0)	4.0 (3.9–4.1)	0.8 (0.7–0.8)			0.9 (0.9–0.9)	0.5 (0.5–0.6)
80–84	1.6 (1.5–1.6)	1.2 (1.2–1.2)	3.7 (3.5–3.8)	0.9 (0.9–1.0)			1.2 (1.2–1.2)	0.8 (0.8–0.8)
⩾85	1.4 (1.3–1.5)	1.4 (1.3–1.4)	3.2 (3.0–3.4)	0.7 (0.6–0.8)			1.7 (1.6–1.7)	1.1 (1.1–1.1)
Females								
Age group (in years)								
0–4								
5–9								
10–14								
15–19	1.4 (0.1–2.6)	0.2 (0.0–0.5)						
20–24	3.3 (2.3–4.2)	0.6 (0.3–0.9)			1.7 (1.1–2.3)	0.6 (0.2–1.0)		
25–29	2.8 (2.2–3.4)	0.5 (0.3–0.7)	2.0 (0.0–4.7)	0.2 (0.0–0.8)	2.1 (2.0–2.4)	0.3 (0.1–0.4)		
30–34	2.1 (1.6–2.5)	0.5 (0.4–0.7)	1.5 (0.6–2.5)	0.2 (0.0–0.5)	1.6 (1.5–1.7)	0.3 (0.2–0.4)		
35–39	2.5 (2.2–2.8)	0.6 (0.5–0.7)	3.1 (2.4–3.7)	0.3 (0.1–0.5)	1.3 (1.2–1.3)	0.3 (0.2–0.3)		
40–44	2.3 (2.1–2.5)	0.5 (0.4–0.6)	3.3 (2.9–3.6)	0.4 (0.2–0.6)	1.0 (0.9–1.0)	0.3 (0.2–0.3)		
45–49	2.2 (2.0–2.3)	0.6 (0.5–0.7)	3.9 (3.6–4.2)	0.5 (0.4–0.7)	0.8 (0.8–0.8)	0.2 (0.2–0.2)		
50–54	2.0 (1.9–2.1)	0.5 (0.4–0.5)	4.2 (4.0–4.4)	0.5 (0.4–0.6)	0.6 (0.6–0.6)	0.2 (0.2–0.3)		
55–59	1.8 (1.7–1.9)	0.5 (0.5–0.5)	4.3 (4.2–4.5)	0.4 (0.4–0.5)	0.5 (0.5–0.5)	0.2 (0.2–0.2)		
60–64	1.7 (1.6–1.8)	0.5 (0.4–0.5)	4.2 (4.1–4.4)	0.5 (0.5–0.6)	0.5 (0.4–0.5)	0.2 (0.2–0.2)		
65–69	1.6 (1.6–1.7)	0.6 (0.5–0.6)	4.0 (3.9–4.1)	0.6 (0.5–0.6)	0.5 (0.5–0.5)	0.3 (0.3–0.3)		
70–74	1.6 (1.5–1.7)	0.7 (0.7–0.7)	4.1 (4.0–4.3)	0.7 (0.7–0.8)	0.5 (0.5–0.5)	0.4 (0.4–0.4)		
75–79	1.5 (1.5–1.6)	0.8 (0.8–0.9)	3.9 (3.8–4.1)	0.9 (0.9–1.0)	0.6 (0.6–0.6)	0.6 (0.6–0.6)		
80–84	1.5 (1.4–1.5)	1.2 (1.1–1.2)	4.3 (4.2–4.6)	1.0 (0.9–1.1)	0.7 (0.6–0.7)	0.9 (0.9–0.9)		
⩾85	1.1 (1.1–1.2)	1.3 (1.2–1.3)	3.8 (3.5–4.1)	0.9 (0.8–1.0)	0.8 (0.8–0.8)	1.2 (1.2–1.2)		

[¶]With 95% confidence intervals. Only observations with at least 20 survivor person-years are included.

Three periods of time after diagnosis were considered in detail: <1 year, between 1 and 5 years and 5–16 years (as only diagnoses from 1990–2006 were available). The first period contained the majority of the cancer-related hospital activity for all cancer survivors, but the period between 1 and 5 years after diagnosis also had a significant amount of hospital activity.

Figure 3 displays the average number of days spent admitted to hospital (per 100 person-days) by cancer survivors in England in 2006, by 5-year age group and broad post-diagnosis period. Only cancer-related episodes of care are presented in this figure. As with all tables and figures, data are omitted if less than 20 person-years of survivorship contributed. Figure 3 shows that, for each age group, the number of days spent admitted to hospital for cancer-related episodes of care (per 100 person-days) was highest in the first year after diagnosis. For colorectal cancer survivors, it was approximately 30 times higher than in the period >5 years after diagnosis. Similarly, for lung cancer survivors, it was approximately 35 times higher than in the period >5 years after diagnosis. The period 1–5 years after diagnosis also contained significantly more cancer-related hospital activity than the period >5 years after diagnosis: it was approximately four times higher among colorectal cancer survivors, and 7 times higher among lung cancer survivors, although Figure 2 shows that much of this increase is due to the high activity in years 2 and 3 after diagnosis. There was a smaller difference in the cancer-related activity in the periods 1–5 and >5 years after diagnosis among prostate and female breast cancer survivors, compared with colorectal and lung cancer survivors.

The number of days spent admitted to hospital for cancer-related episodes of care (per 100 person-days) in the first year after diagnosis was significantly higher among survivors over the age of 75 years, compared with younger survivors. This is most noticeable for survivors of prostate cancer – those aged ⩾85 years spent around three times as many days admitted in the first year after diagnosis, compared with those aged 70–74 years.

Table 2 Mean admitted patient hospital activity (cancer related and non-cancer related) among cancer survivors, England, 2006. Number of days in hospital per 100 person-days,[¶] by time since diagnosis

	Cancer diagnosis							
	Colon, rectum and anus		Lung, bronchus and trachea		Breast		Prostate	
	Cancer related	Non-cancer related	Cancer related	Non-cancer related	Cancer related	Non-cancer related	Cancer related	Non-cancer related
Males								
Years since diagnosis								
<1	7.7 (7.7–7.8)	1.3 (1.3–1.4)	9.7 (9.6–9.7)	1.0 (1.0–1.1)			1.8 (1.8–1.9)	0.5 (0.5–0.5)
1–2	1.5 (1.4–1.5)	1.0 (1.0–1.1)	3.0 (2.9–3.1)	0.6 (0.5–0.7)			0.8 (0.8–0.9)	0.4 (0.4–0.5)
2–3	1.0 (1.0–1.1)	0.9 (0.9–0.9)	1.7 (1.5–1.8)	0.7 (0.7–0.8)			0.7 (0.7–0.8)	0.5 (0.5–0.5)
3–4	0.8 (0.8–0.9)	0.9 (0.8–0.9)	1.0 (0.9–1.2)	0.6 (0.5–0.7)			0.7 (0.7–0.8)	0.5 (0.5–0.6)
4–5	0.6 (0.6–0.7)	0.9 (0.8–0.9)	0.7 (0.6–0.9)	0.7 (0.6–0.8)			0.7 (0.7–0.8)	0.5 (0.5–0.6)
5–6	0.5 (0.4–0.6)	1.0 (0.9–1.0)	0.5 (0.3–0.7)	0.7 (0.6–0.8)			0.7 (0.7–0.8)	0.6 (0.6–0.7)
6–7	0.4 (0.4–0.5)	0.9 (0.9–1.0)	0.5 (0.3–0.7)	0.8 (0.6–0.9)			0.7 (0.7–0.8)	0.6 (0.6–0.7)
7–8	0.4 (0.3–0.4)	0.9 (0.9–1.0)	0.4 (0.2–0.7)	0.7 (0.5–0.8)			0.8 (0.8–0.9)	0.6 (0.6–0.7)
8–9	0.3 (0.2–0.4)	0.8 (0.7–0.9)	0.3 (0.1–0.5)	0.7 (0.6–0.8)			0.8 (0.7–0.8)	0.8 (0.7–0.8)
9–10	0.2 (0.2–0.3)	0.6 (0.6–0.7)	0.2 (0.0–0.4)	0.4 (0.3–0.5)			0.8 (0.8–0.9)	0.6 (0.6–0.7)
10–11	0.3 (0.2–0.4)	0.6 (0.5–0.6)	0.3 (0.1–0.5)	0.2 (0.1–0.3)			0.8 (0.7–0.9)	0.6 (0.6–0.7)
11–12	0.3 (0.2–0.4)	0.5 (0.5–0.6)	0.2 (0.0–0.4)	0.3 (0.2–0.5)			0.8 (0.7–0.9)	0.6 (0.5–0.7)
12–13	0.2 (0.1–0.3)	0.4 (0.3–0.5)	0.2 (0.0–0.4)	0.2 (0.1–0.4)			0.7 (0.6–0.8)	0.5 (0.5–0.6)
13–14	0.3 (0.2–0.4)	0.4 (0.3–0.5)	0.1 (0.0–0.3)	0.2 (0.0–0.4)			0.7 (0.6–0.9)	0.5 (0.4–0.6)
14–15	0.2 (0.1–0.3)	0.4 (0.3–0.5)	0.2 (0.0–0.4)	0.3 (0.1–0.5)			0.6 (0.5–0.8)	0.5 (0.4–0.7)
15–16	0.2 (0.1–0.4)	0.5 (0.4–0.6)	0.2 (0.0–0.5)	0.1 (0.0–0.3)			0.6 (0.4–0.8)	0.5 (0.4–0.6)
Females								
Years since diagnosis								
<1	7.8 (7.7–7.9)	1.3 (1.3–1.4)	9.7 (9.6–9.8)	1.0 (0.9–1.0)	2.6 (2.5–2.6)	0.5 (0.5–0.5)		
1–2	1.3 (1.3–1.4)	1.0 (1.0–1.0)	2.7 (2.6–2.9)	0.6 (0.6–0.7)	0.6 (0.5–0.6)	0.5 (0.5–0.5)		
2–3	0.9 (0.8–0.9)	1.0 (0.9–1.0)	1.7 (1.5–1.8)	0.6 (0.5–0.7)	0.5 (0.4–0.5)	0.5 (0.5–0.5)		
3–4	0.7 (0.6–0.8)	0.8 (0.8–0.9)	1.1 (0.9–1.3)	0.6 (0.5–0.8)	0.4 (0.4–0.5)	0.5 (0.5–0.5)		
4–5	0.5 (0.4–0.6)	0.8 (0.8–0.9)	0.7 (0.5–0.9)	0.8 (0.7–1.0)	0.4 (0.4–0.4)	0.5 (0.5–0.5)		
5–6	0.4 (0.3–0.5)	0.9 (0.9–0.9)	0.8 (0.5–1.1)	0.7 (0.6–0.8)	0.3 (0.3–0.4)	0.5 (0.5–0.5)		
6–7	0.3 (0.2–0.4)	0.9 (0.8–0.9)	0.3 (0.1–0.6)	0.7 (0.5–0.9)	0.3 (0.3–0.4)	0.5 (0.5–0.6)		
7–8	0.2 (0.1–0.3)	0.9 (0.9–1.0)	0.3 (0.1–0.5)	1.1 (0.9–1.3)	0.3 (0.3–0.4)	0.5 (0.5–0.5)		
8–9	0.3 (0.2–0.4)	1.0 (1.0–1.1)	0.3 (0.0–0.6)	0.5 (0.3–0.7)	0.3 (0.3–0.3)	0.5 (0.4–0.5)		
9–10	0.2 (0.1–0.3)	0.6 (0.6–0.7)	0.2 (0.0–0.5)	0.5 (0.3–0.8)	0.3 (0.3–0.3)	0.3 (0.3–0.3)		
10–11	0.2 (0.1–0.3)	0.4 (0.4–0.5)	0.2 (0.0–0.4)	0.2 (0.0–0.4)	0.3 (0.2–0.3)	0.2 (0.2–0.2)		
11–12	0.2 (0.1–0.3)	0.3 (0.3–0.4)	0.3 (0.0–0.6)	0.1 (0.0–0.3)	0.3 (0.2–0.3)	0.2 (0.2–0.3)		
12–13	0.2 (0.1–0.3)	0.3 (0.3–0.4)	0.3 (0.0–0.7)	0.3 (0.1–0.5)	0.3 (0.2–0.3)	0.2 (0.2–0.2)		
13–14	0.2 (0.1–0.3)	0.3 (0.3–0.4)	0.2 (0.0–0.4)	0.1 (0.0–0.3)	0.3 (0.3–0.4)	0.2 (0.2–0.3)		
14–15	0.1 (0.0–0.2)	0.4 (0.3–0.4)	0.2 (0.0–0.7)	0.1 (0.0–0.3)	0.2 (0.2–0.3)	0.2 (0.2–0.3)		
15–16	0.1 (0.0–0.3)	0.3 (0.2–0.4)	0.2 (0.0–0.5)	0.1 (0.0–0.4)	0.3 (0.2–0.3)	0.2 (0.2–0.3)		

[¶]With 95% confidence intervals.

DISCUSSION

This paper presents a person-time analysis of a linked cancer registry and hospital activity data set for England. Age and time since diagnosis were anticipated to be important factors in the level of health-care utilisation among cancer survivors, and were studied in detail. The linked data sets allowed an analysis of all recorded episodes of hospital inpatient or day case health care in England during 2006 among registered cancer survivors, with at least one cancer-related episode recorded in the HES in the period 1990–2006. A person-time approach was used throughout, which made it possible to quantify the intensity of hospitalisation, as well as the associated burden on cancer survivors and the health service, by considering the amount of time spent in hospital, rather than just the number of admissions. Person-time of survivorship and person-time of hospitalisation were classified according to the variables of interest (age and time since diagnosis) – this made it possible to describe in detail the ways in which acute health-service utilisation among survivors was related to these factors.

It is not possible to distinguish between the cancer survivors who had no recorded hospital activity and those for whom the linkage process between the cancer registry and HES data failed. It was therefore not possible to exclude patients with 'no match' to HES. This results in an underestimation of the amount of hospital activity among the cancer survivors, but this is believed to be only a small effect.

Each record in the HES data had up to 14 diagnostic codes (using the ICD-10 classification), and 12 operation procedure codes (using the OPCS4 classification), with the first of each of these codes intended to indicate the primary diagnosis/intent of the episode (Maddams *et al*, 2011). For simplicity, episodes of care were broadly categorised as 'related to cancer' or 'not related to cancer' according to the diagnostic codes only – an episode was considered to be cancer related if any of the diagnostic codes was between C00 and C97 (excluding C44). This was an intentionally broad definition, designed to negate any possible regional variation in coding of primary diagnosis in the HES data, and to take account of the wide range of health problems associated with cancer and the side effects of its treatment. Defined thus, 'cancer-related' hospitalisation for a given survivor may be considered as that directly or indirectly caused by, or associated with, a cancer diagnosis. A future area of study may be to define 'cancer-related' HES admissions according to the clinical procedure codes, as well as the admission codes.

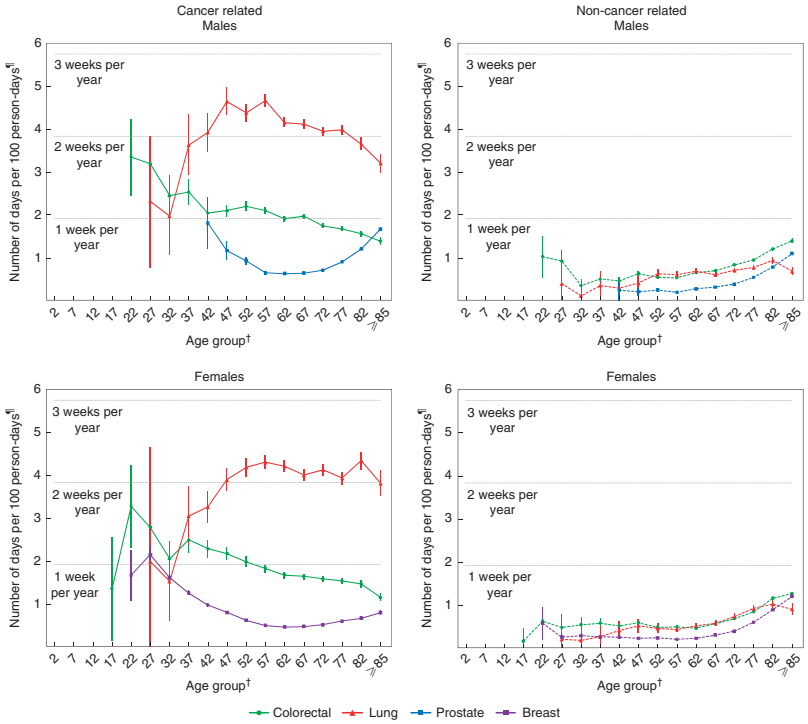


Figure 1 Mean admitted patient hospital activity among cancer survivors, England, 2006. Number of days admitted to hospital per 100 person-days¶, by 5-year age group. †Centre of 5-year age group. ¶With 95% CIs. Only data points with at least 20 survivor person-years are plotted.

Perhaps a more precise way of defining the health-care burden directly attributable to cancer would be to compare hospitalisation of any kind among cancer survivors with that observed in a randomly selected age- and sex-matched subset of the cancer-free population (or, more practically, the general population). This would allow some measure of the 'background' hospitalisation experienced by the cancer-free or general population to be removed from the levels observed in cancer survivors. However, no such data set was available to the authors at the time of analysis – the Cancer Registries' National HES extract only contains episodes of care for patients with at least one episode 'for or with cancer' – and thus it was only possible to make the broad cancer/non-cancer-related distinction at a population level. It is acknowledged, however, that at an individual level it may not be possible to make such a clear distinction between cancer-related and non-cancer-related health-service utilisation. Non-cancer-related acute health-care utilisation was shown to generally increase with age, but remained roughly constant regardless of time since diagnosis, indicating that it may be a good minimum measure of 'background' acute health-care utilisation. That said, there is higher non-cancer-related acute health-care utilisation in the first year following diagnosis, particularly among those with lung and colorectal cancer. This may be explained by the fact that some survivors may have had an emergency admission to hospital during which they were diagnosed with cancer. In addition, some genuine cancer-related hospital activity may not be coded as such in the HES.

The primary motivation for using a person-time approach in this analysis was the need to properly account for the length of time cancer survivors spent 'at risk' of hospitalisation during the period of analysis, and how this depended on both age and time since diagnosis. The person-time methods allowed a precise assessment of the influence of these variables on the amount of hospitalisation among cancer survivors. The number of days survivors spent admitted to hospital (per 100 person-days) provides a more revealing assessment of the burden of cancer on the health service than a count of the number of admissions. However, the limitation of this approach is that, as person-time is pooled for all survivors in the population, it provides only a measure of the mean health-service utilisation in the population of survivors and obscures the underlying distribution. For example, on average male survivors of colorectal cancer who were no more than 1 year from diagnosis in 2006 spent 9 days per 100 person-days admitted to hospital (equivalent to approximately one month per year); however, in reality, many such survivors spent no time admitted to hospital at all, whereas others spent more than 9 days per 100 as an admitted hospital patient. This variation in the intensity of hospitalisation across survivors of a similar age and post-diagnosis period is considered in Maddams *et al* (2011), which used a population person-time analysis to estimate the

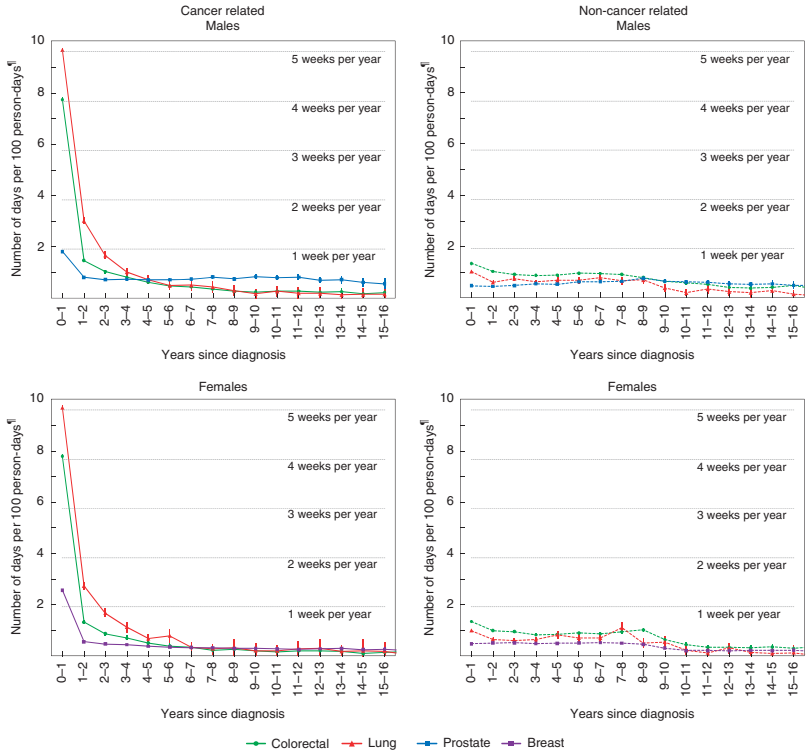


Figure 2 Mean admitted patient hospital activity among cancer survivors, England, 2006. Number of days admitted to hospital per 100 person-days¶, by 1-year since diagnosis period. ¶With 95% CIs. Only data points with at least 20 survivor person-years are plotted.

proportion of survivors with a low or high level of health-service utilisation, as well as those with none. It was shown that the underlying distribution of hospital activity is skewed towards those with no hospital activity, particularly in the period more than 1 year after diagnosis, and it should therefore be kept in mind that the mean number of days spent in hospital per 100 person days, as described in this paper, is largely influenced by the small number of survivors who have a very high level of hospital activity.

Another limitation of this approach when used to assess the burden of cancer on the health service is that each day of hospitalisation is treated equally and the variation in the burden presented by different types of admission and procedures is not taken into account. For example, some survivors will be admitted to hospital for routine observation and monitoring, whereas others will be admitted for complicated operations that consume large amounts of hospital resources and require intense periods of rehabilitation.

The majority of cancer-related health-service utilisation occurred during the first year following diagnosis, for all types of cancer studied. Most cancer patients receive some form of care or treatment as soon as possible after diagnosis, and thus it is perhaps unsurprising that this period contains a large amount of hospital activity. However, this study also shows that there is a significant amount of cancer-related health-service utilisation in the period 1–5 years after diagnosis, particularly among survivors of lung and colorectal cancers. This is no longer the initial treatment phase, but is indicative of the ongoing needs of cancer survivors.

The highest levels of cancer-related acute health-service utilisation were observed in survivors of the relatively poor prognosis cancers, but these differences largely disappeared more than 5 years after diagnosis. This indicates that the worse prognosis cancers required more intensive treatment regimes (including end-of-life care) in the short to medium term after diagnosis, but not in the long term.

Prostate cancer survivors, despite having lower levels of cancer-related acute health-service utilisation in the first year after diagnosis than those with colorectal, lung or breast cancer, actually had the highest levels five or more years after diagnosis (Figure 2). This effect is the result of the relatively high levels of hospitalisation among prostate cancer survivors aged over 70 years and five or more years from diagnosis (Figure 3). A more detailed investigation (not included here) showed that these survivors experienced a large amount of hospital activity recorded in the HES data with a non-cancer primary diagnosis code, but with prostate cancer recorded as one of the supplementary diagnoses, and it was largely these episodes that resulted in the relatively high number of days spent admitted to hospital (per 100 person-days)

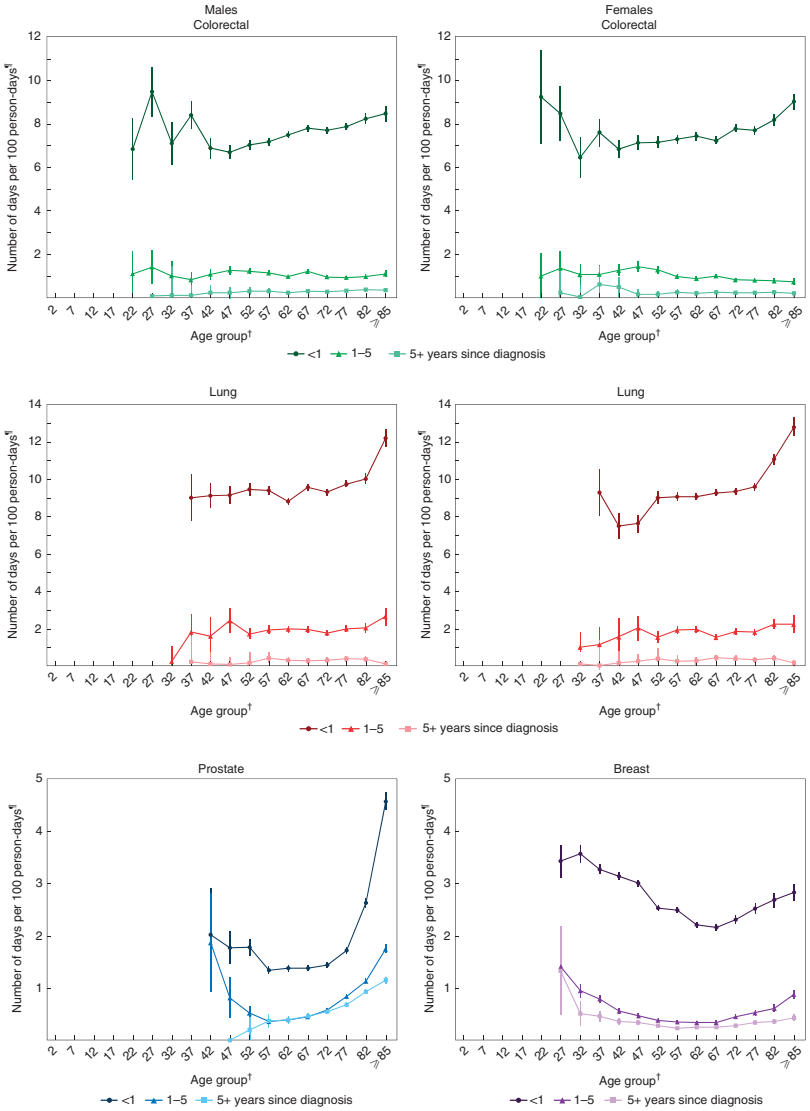


Figure 3 Mean cancer-related admitted patient hospital activity among cancer survivors, England, 2006. Number of days admitted to hospital per 100 person-days¶, by 5-year age group and broad time since diagnosis period. ¶With 95% CIs. Only data points with at least 20 survivor person-years are plotted. †Centre of 5-year age group.

among this group, compared with equivalent survivors of other cancers. In the analysis, these episodes were considered to be 'cancer related' because cancer was recorded as one of the supplementary diagnosis codes in the HES data. However, the recorded primary diagnosis codes, together with the operation procedure codes, indicate that these episodes were not generally related directly to prostate cancer. For example, the most common procedure codes were endoscopic examinations of the bladder and urethral catheterisations of the bladder, and the most common primary diagnosis codes were for disorders of the urinary system.

There are several physiological changes that occur in men as they get older and lead to alterations in lower urinary tract function, making urinary disorders in elderly men common (Dubeau, 2006; Griebling, 2008). Many prostate cancer survivors are likely to be closely monitored in urology clinics for many years after initial diagnosis – often an extended period of either 'active surveillance' or 'watchful waiting' is pursued, especially if the cancer is less aggressive and treatment is not immediately necessary (Cancer Research UK, 2010a). This may explain the greater hospital activity recorded among prostate cancer survivors over the age of 70 years many years after diagnosis, compared with male lung or colorectal cancer survivors – it is possible that the increased observation and monitoring of these survivors leads to a high level of urological intervention, which is only indirectly related to their prostate cancer diagnosis.

The patterns of cancer-related health-service utilisation by age, as described in Figure 1, must be viewed in the context of the findings presented in Figure 3. For example, older cancer survivors are more likely to be more than 1 year from diagnosis, compared with younger survivors, and this alone explains the apparent decreasing levels of utilisation as age increases among lung and colorectal cancer survivors. Acute health-service utilisation in the first year after diagnosis was generally highest in survivors aged over 70 years. Initial cancer treatment can be physically very arduous, particularly for older patients who may be more frail and suffering from comorbidities, resulting in more frequent and extended admissions to hospital. Particularly striking was the threefold increase in cancer-related acute health-service utilisation in the first year after diagnosis among prostate cancer survivors aged $\geqslant$85 years compared with those aged 65–69 years (Figure 3). Similar, but not as significant, increases were observed among survivors of colorectal, lung and breast cancers. Since the early 1990s, the PSA test has increasingly been available (usually to men aged over 50 years of age) as a screening tool for prostate cancer, although in the UK no organised PSA screening programme is in place. This test is still controversial and considered to potentially result in overdiagnosis and overtreatment of prostate cancers that would otherwise never have become symptomatic (Barry, 2009; van Leeuwen *et al*, 2010). Recorded incidence and survival have, accordingly, greatly increased since the test's introduction (Evans and Møller, 2003; Cancer Research UK, 2010b, c), as have the number of prostate cancer survivors (Maddams *et al*, 2009). Men diagnosed with prostate cancer over the age of 70 years are more likely to have symptomatic disease requiring intensive initial treatment compared with those aged 50–70 years who are more likely to have been diagnosed via the PSA test. Similarly, the observed lower acute health-service utilisation in recently diagnosed breast cancer survivors aged 50–70 years (the age range in which women are routinely invited to attend breast screening units in England (NHS Breast Cancer Screening Programme, 2011), compared with other age groups, reflects some of the benefits of early detection offered by the screening programmes.

This paper highlights the significant effect that time since diagnosis has on the average amount of hospitalisation experienced by cancer survivors. The majority of cancer-related admitted hospital episodes of care occurred in the first year following diagnosis (when initial cancer treatment takes place); however, there was also a significant amount of hospital activity in the period 1–5 years after diagnosis, particularly among survivors of colorectal and lung cancers, which is indicative of the ongoing consequences of cancer and its treatment. Other work has shown that the final year before death also contains significant amounts of hospitalisation of cancer survivors (Maddams *et al*, 2011). These findings help to understand the burden of cancer on the health service, but further work is still required to identify potential areas of unmet needs among cancer survivors.

ACKNOWLEDGEMENTS

This paper is a contribution from the National Cancer Intelligence Network and is based on the information collected and quality assured by the regional cancer registries in UK (www.ukacr.org; www.ncin.org.uk). The Thames Cancer Registry in King's College London receives funding from the Department of Health (DoH) for England. The views expressed in the article are those of the authors and not necessarily those of the DoH. The Clinical Operational Research Unit receives funding from the DoH for England Policy Research Programme. This article is sponsored by Macmillan Cancer Support.

REFERENCES

Barry MJ (2009) Screening for prostate cancer – the controversy that refuses to die. *N Engl J Med* **360**(13): 1351–1354

Budescu DV (1985) Analysis of dichotomous variables in the presence of serial dependence. *Psychol Bull* **97**(3): 547–561

Cancer Research UK (2010a) Monitoring prostate cancer. http://www.cancerhelp.org.uk/type/prostate-cancer/treatment/monitoring-prostate-cancer (accessed 23 June 2011)

Cancer Research UK (2010b) Prostate cancer – survival statistics. http://info.cancerresearchuk.org/cancerstats/types/prostate/survival/ (accessed 23 June 2011)

Cancer Research UK (2010c) Prostate cancer – UK incidence statistics. http://info.cancerresearchuk.org/cancerstats/types/prostate/incidence/ (accessed 23 June 2011)

Carstensen B (2007) Examples and programs for splitting follow-up time in cohort studies. http://staff.pubhealth.ku.dk/~bxc/Lexis/ (accessed 23 June 2011)

Dubeau CE (2006) The aging lower urinary tract. *J Urol* **175**(3 Pt 2): S11–S15

Evans HS, Møller H (2003) Recent trends in prostate cancer incidence and mortality in southeast England. *Eur Urol* **43**(4): 337–341

Griebling T (2008) Urinary incontinence and voiding dysfunction in elderly men. *Curr Bladder Dysfunct Rep* **3**(4): 241–246

Klotz J (1973) Statistical inference in Bernoulli trials with dependence. *Ann Statist* **1**(2): 373–379

Maddams J, Brewster D, Gavin A, Steward J, Elliott J, Utley M, Moller H (2009) Cancer prevalence in the United Kingdom: estimates for 2008. *Br J Cancer* **101**(3): 541–547

Maddams J, Utley M, Moller H (2011) Levels of acute health service use among cancer survivors in the United Kingdom. *Eur J Cancer* **47**(14): 2211–2220

NHS Breast Cancer Screening Programme (2011) About breast screening. http://www.cancerscreening.nhs.uk/breastscreen/about-breast-screening.html (accessed 23 June 2011)

NHS Health and Social Care Information Centre (2010) HES User Guide. http://www.hesonline.nhs.uk/Ease/servlet/ContentServer?siteID=19378&categoryID=87 (accessed 23 June 2011)

van Leeuwen PJ, Connolly D, Tammela TLJ, Auvinen A, Kranse R, Roobol MJ, Schroder FH, Gavin A (2010) Balancing the harms and benefits of early detection of prostate cancer. *Cancer* **116**(20): 4857–4865